Service Function Chaining: Design Considerations, Analysis & Recommendations
draft-boucadair-sfc-design-analysis-01

Abstract

   This document aims at analyzing the various design options and
   providing a set of recommendations for the design of Service Function
   Chaining solution(s). Note:

   o The analysis does not claim to be exhaustive. The list includes a
      preliminary set of potential solutions; other proposals can be
      added to the analysis if required.

   o The analysis is still ongoing. The analysis text will be updated
      to integrate received comments and inputs.

   o Sketched recommendations are not frozen. These recommendations
      are provided as proposals to kick-off the discussion and to
      challenge them.

   o The analysis does not cover any application-specific solution
      (e.g., HTTP header) because of the potential issues inherent to
      (TLS) encrypted traffic.

   o The analysis will be updated to take into account the full set of
      SFC requirements.

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 RFC 2119 [RFC2119].

Status of This Memo

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

Table of Contents

1. Introduction

   This document aims at analyzing the various design options and
   providing a set of recommendations for the design of Service Function
   Chaining solution(s). The conclusions of this analysis, once stable,
   will be recorded in the framework document.

   The overall problem space is described in
   [I-D.quinn-sfc-problem-statement]. A list of requirements is
   available at [I-D.boucadair-sfc-requirements].

2. Terminology

   The reader should be familiar with the terms defined in
   [I-D.boucadair-sfc-framework].

3. Scope

   This document identifies potential solutions to fulfill the design
   requirements documented in [I-D.boucadair-sfc-requirements].
   Particularly, it focuses on the following design objectives:

   1. Which information to include in the SFC header? (see Section 4)

   2. How to mark packets to indicate they belong to a given Service
       Function Chain (SFC) (see Section 5) and in which channel the SFC
       header is to be conveyed (see Section 6)?

   3. How to select a differentiated set of policies at a given Service
       Function (SF)? (see Section 5)

   4. How to select the forwarding path of a given flow that needs to
       be processed according to a set of Service Functions which must
       be invoked in a given order? (see Section 7)

   Other design issues will be documented in future versions if
   required.

4. Service Function Chaining Header

   This section identifies the main design points to be agreed upon so
   as to guide the forthcoming specification effort of the Service
   Function Chaining Header.

4.1. Why a Subscriber Identifier Does Not Need to be part of the Header?

   Injecting (or leaking a Subscriber Identifier (subscriber-ID)) in the
   Service Function Chaining Header is not motivated by current
   deployment practices which consider that per-subscriber policies are
   enforced by only a few nodes (especially in the access network
   segment). As such, the service chaining solution does not require
   per-subscriber policies to be supported by all involved (SF) nodes.

   Moreover, the enforcement of some policies can be driven by a subset
   of the information contained in the packets (e.g., source IP address,
   IPv6 prefix, etc.). Conveying the SF Map Index is sufficient to
   guide a Service Function to select the set of policies to be enforced
   for a given packet that belongs to a flow.

   Note, current deployments may require an explicit subscriber-ID only
   during the authentication and authorization phases. These
   deployments do not require explicit subscriber-ID information to be
   conveyed when sending actual data traffic.

   The Service Chaining approach does not aim to change how
   authorization and per-subscriber policies are enforced. As a
   consequence, explicit Subscriber-ID is not mandatory to be included
   in the Service Chaining Header.

   In some contexts, the subscriber-ID is not available to the node that
   is responsible for enforcing per-subscriber policies. A typical
   example where complications may arise is a deployment context
   characterized as follows:

   o Overlapping IP address pools are in use.

   o NAT function is not collocated with the GGSN or BNG.

Enforcing for instance per-subscriber port quota requires an
additional information to uniquely disambiguate hosts having the same
address (called HOST_ID, [RFC6967]). This problem is not specific to
the Service Function Chaining, but it is encountered in many other
use cases ([I-D.boucadair-intarea-host-identifier-scenarios]).

Within the context of SFC, two solutions can be adopted:

1. Implement a solution similar to what is specified in [RFC6674].
    This means that the subscriber-ID is passed only to the node that
    enforces the per-subscriber policies without leaking it to other
    downstream SFs. In such case, the node that inserts the
    subscriber-ID is not part of the SFC-enabled domain. This
    solution does not require the insertion of the subscriber-ID in
    the SFC header.

2. Define a subscriber-ID optional field in the SFC header. This
    optional field can be defined as an optional 64-bit field to
    accommodate the mobile case (e.g., inject an IMSI (International
    Mobile Subscriber Identity) identifier as a subscriber-ID).

```
+---------------------------------------------------------------------+
| Proposed Recommendation                                             |
+---------------------------------------------------------------------+
| It is NOT RECOMMENDED to encode a subscriber-ID                     |
| as a mandatory field of the SFC header.                             |
+---------------------------------------------------------------------+
```

4.2. Fixed vs. Variable Length of the SFC Map Index

The number of Service Chains to be instantiated is deployment-
specific. It depends on the business context and engineering
practices that are internal to each administrative entity. To ensure
a better flexibility as a function of the service chains that are
theoretically supported, a first design consideration is to decide
whether there is a need for a fixed field or a variable length field.

A field with a variable length is flexible enough to accommodate as
many Service Function Chains as required for each deployment context.
An administrative entity will need to tweak the length of this field
to meet its own deployment requirements (e.g., set the length in all
involved nodes to 8 bits, 16 bits, 32 bits or even more).

A field with a fixed length would lead to a better performance
(mainly because of a simplified processing).

4.3. Recommended Length

An 8-bit field would be sufficient to accommodate deployment contexts
that assume a reasonable set of SF Maps. A 16-bit field would be
more flexible and would allow to enable large service chains (e.g.,
to accommodate the requirement discussed in
[I-D.boucadair-sfc-requirements]). A 32-bit field would fulfill the
needs for deployments with very large Service Function Chains.

```
+---------------------------------------------------------------------+
| Proposed Recommendation                                             |
+---------------------------------------------------------------------+
| It is RECOMMENDED to use a 32-bit field to encode                   |
| the SF Map Index                                                    |
+---------------------------------------------------------------------+
```

4.4. Extensibility

The header can be extended in the future, based on the experience
that will be gained during operational deployments. As such, the
header does not need to include any protocol version field nor any
reserved bits to disambiguate between two variants of the header.

Implementations supporting the service chaining solution can be
upgraded following current best practices in the field.

```
+---------------------------------------------------------------------+
| Proposed Recommendation                                             |
+---------------------------------------------------------------------+
| It is NOT RECOMMENDED to reserve bits to anticipate future          |
| extension needs. Backward compatibility between two versions of     |
| the header can be ensured by consistent system                      |
| setup & configuration.                                              |
+---------------------------------------------------------------------+
```

5. Format of the Service Function Chaining Header

This section proposes and discusses some formats to encode the
Service Chaining Header. An analysis is also included in this
section.

   [NOTE: Other proposals may be added to this section.]

5.1. Format 1: Single Marking Code Point

The RBNF format [RFC5511] of the header is shown in Figure 1:

```
                    <SFC Header> ::=  <SF Map Index>
```

                       Figure 1: Single Marking Code Point

   This format is characterized as follows:

   o A policy table is provisioned to each SF Node. This policy table
      includes the locator(s) of the possible SF next hop and the SF Map
      Index list to help detect any Service Function Loop.

   o Classifiers are provisioned with classification rules to decide
      which code point to use for a received packet.

   o Fragmentation risk is minimized because the header is compacted.

   o Multiple profiles can be supported per SF Node; each profile is
      identified with a Service Function Identifier.

   o The classifier behavior is simplified.

   o Separating the policies channel from the marking behavior prevents
      potential DDoS (e.g., common to any source routing scheme.)

   o The lookup in the SFC Policy Table is not a concern because it is
      not expected to provision SFC Policy Tables with an amount of
      information (e.g., like the size of the global routing table).

5.2. Format 2: Marking Code Point & Profile Index

   The RBNF format of the header is shown in Figure 2:

```
        <SFC Header> ::=  <SF Map Index>
                         <Service Function Map>

        <Service Function Map> ::= <Service Function> ...

        <Service Function> ::=  <Service Function Identifier>
                                <Profile Identifier>
```

              Figure 2: Marking Code Point & Profile Index

   This format is characterized as follows:

   o The list of SF Locator(s) is provisioned out of band to each SF
      Node.

   o Classifiers are provisioned with classification rules to decide
      which code point is to be used for a received packet.

o Fragmentation risks are not minimized.

o The classifier needs to be configured with a list of profiles/
   contexts per Service Function.

o The classifier behavior is not simplified since it must also
   encode in each incoming packet the full list of functions to be
   performed by each Service Function hop.

5.3. Format 3: Explicit Route List

The RBNF format of the header is shown in Figure 3:

```
<SFC Header> ::=  <Total number of Service Function hops>
                  <Current hop Index>
                  <Service Function Map>

<Service Function Map> ::= <Service Function> ...

<Service Function> ::=  <IP ADDRESS>
                        <Profile Identifier>
```

                     Figure 3: Explicit Route List

The procedure at a non-reclassifying node is to validate that the IP
address of the SF at the current index matches one of the SF's own IP
addresses and then to find the profile identifier by its indicated
identifier. Once the local Service Function is invoked, if the
packet needs to be forwarded to the next Service Function hop, the
local node simply increments the current hop index and rewrites the
outer IP header with the next hop's IP address.

This format is characterized as follows:

o Classifiers are provisioned with classification rules to decide
   which code point is to be used for a received packet.

o Fragmentation risks are not minimized.

o The classifier needs to be configured with a list of profiles/
   contexts per Service Function.

o The classifier is also responsible for load balancing. This makes
   the classifier more complex.

o The classifier behavior is not simplified since it must also
   encode in each incoming packet the full list of policies to be
   performed by each Service Function node.

5.4. Format 4: Compact Explicit Route List

A variant of the previous format is depicted in the RBNF format of
the header shown in Figure 4. Instead of including the explicit
route list (Figure 3), IP addresses of SFs are configured out of band
but each of these addresses is identified with a unique identifier.
These identifiers are indicated in the Service Chaining Header.

```
     <SFC Header> ::=  <Total number of Service Function hops>
                       <Current hop Index>
                       <Service Function List>

     <Service Function List> ::= <Service Function> ...

     <Service Function> ::=  <IP Address ID>
                             <Profile Identifier>
```

                  Figure 4: Compact Explicit Route List

This proposal suffers from the same drawbacks as the previous format.

5.5. Analysis

Given the design motto that says:

   "A protocol design is complete not when you can't think of any
   more things to add, but when you have removed everything you can
   and you can't see how to remove any more",

the proposed format must be as simple as possible while meeting the
requirements discussed in [I-D.boucadair-sfc-requirements]. The
simplicity argument is further discussed in [RFC3439] and [Robust].

Based on the above analysis, the proposal that is simple, minimizes
fragmentation, optimizes the behavior of the classifier and SF Nodes,
and that prevents potential DDoS attacks is the one discussed in
Section 5.1.

6. Where To Convey the Chaining Marking Information In A Packet?

This section lists a set of candidate solutions to convey the Service
Chaining Header.

6.1. Use IPv6 Flow Label

   The use of the 20-bit Flow Label field in the IPv6 header [RFC6437]
   can be considered as a candidate solution to convey the SF Map Index.

   The following comments can be made for this candidate solution:

   o This proposal requires all packets are transported over IPv6.
      This should not be considered as a limitation for some
      deployments.
   o Intermediate Nodes must not alter the content of the Flow Label
      field.
   o This proposal can apply to any transport protocol.
   o The use of the IPv6 Flow Label may interfere with other usages of
      the flow label such as Equal Cost Multipath (ECMP) or Link
      Aggregation (LAG) [RFC6438]. The Flow Label bits need to be
      combined at least with bits from other sources within the packet,
      so as to produce a constant hash value for each flow and a
      suitable distribution of hash values across flows [RFC6437].
   o A 20-bit field to convey the SF Map Index allows to enable Service
      Function Chains of a large size range.
   o This proposal does not allow to convey additional information than
      the SF Map Index (if needed).
   o The Flow Label is present in all fragments, SF Nodes do not need
      to maintain any state to handle a fragmented packet.
   o Altering the value of the Flow Label field does not interfere with
      the use of IPsec [RFC6438].
   o Carrying the SF Map Index in the IPv6 Flow Label allows to:

      * De-correlate packet marking from forwarding constraints.
      * Avoid requiring an internal tagging mechanism to each SF Node
         to preserve the same marking in the outgoing interface as the
         one received in an incoming interface.

```
   +---------------------------------------------------------------------+
   | Proposed Recommendation                                             |
   +---------------------------------------------------------------------+
   | It is tempting to use the Flow Label, but the 20-bit length of      |
   | the Flow Label field is conflicting with the recommended 32-bit     |
   | length discussed in Section 4.3.                                    |
   |                                                                     |
   | The use of Flow Label is NOT RECOMMENDED.                           |
   +---------------------------------------------------------------------+
```

6.2. Use the DS Field

Another alternative to convey the SF Map Index is to use the
Differentiated Services (DS) field [RFC2474] [RFC2475] (for both IPv4
and IPv6).

The following comments can be made for this proposal:

o This proposal overloads the semantics of the DS field.
o Having 64 possible values may not accommodate deployments with a
   large number of service chains (see Section 4.3).
o This proposal can apply to any transport protocol.
o The use of the DS field for service chaining purposes may
   interfere with other usages such as Traffic Engineering (TE) or
   Quality of Service (QoS).

   * This issue can be mitigated by fragmenting the DS space into to
      distinct set of values; each set dedicated for a specific
      usage. An administrative entity can use the first bits for
      service chaining and other remaining bits for QoS for instance.
   * Splitting the DS space reduces the number of possible service
      chains to be configured per administrative domain.

```
+---------------------------------------------------------------------+
| Proposed Recommendation                                             |
+---------------------------------------------------------------------+
| The use of DS field is NOT RECOMMENDED.                             |
+---------------------------------------------------------------------+
```

6.3. Use IP Identification Field

The IPv4 ID (Identification field of IP header, i.e., IP-ID) can be
used to insert the SF Map Index. The classifier rewrites the IP-ID
field to insert the SF Map Index (16 bits). The classifier must
follow the rules defined in [RFC6864]; in particular, the same SF Map
Index is not reassigned during a given time interval. Note:

o This usage is not consistent with the fragment reassembly use of
   the Identification field [RFC0791] or the updated handling rules
   for the Identification field [RFC6864].
o Complications may arise if the packet is fragmented before
   reaching the Classifier. To appropriately handle those packet
   fragments, the classifier will need to maintain a lot of state.
o Preserving the same value when crossing all intermediate SFs may
   be difficult (e.g., an invoked SF can be a NAT).
o This proposal assumes packets are transported over IPv4 (plain or
   encapsulated mode). This may not be considered as a limitation
   for some deployments.

```
+----------------------------------------------------------------------+
| Proposed Recommendation                                              |
+----------------------------------------------------------------------+
| Using the IP-ID as a channel to convey the SF Map Index is NOT       |
| RECOMMENDED.                                                         |
+----------------------------------------------------------------------+
```

[6.4](). Use IPv4 SSRR/LSRR Option

Another candidate channel to convey the Service Chaining Header is to
use the IPv4 SSRR/LSRR options [[RFC0791]()]. These options can be
inserted by the classifier following the pre-configured
classification rules. Note:

o Some general recommendations documented in
   [[I-D.ietf-opsec-ip-options-filtering]()] and [[RFC6192]()] need to be
   taken into account.
o This proposal assumes packets are transported over IPv4 (plain or
   encapsulated mode). This may not be considered as a limitation
   for some deployments.
o This proposal can apply to any transport protocol.
o Encoding the full list of intermediate SF Nodes will exacerbate
   fragmentation issues.
o Injecting an additional IP option by the classifier introduces
   some implementation complexity in the following cases: The packet
   has the MTU size (or is close to it), and the option space is
   exhausted.
o Legacy nodes must be configured to not strip this option.
o Processing the IP option may degrade the performance of involved
   SF nodes.

```
+----------------------------------------------------------------------+
| Proposed Recommendation                                              |
+----------------------------------------------------------------------+
| Using the IPv4 SSRR/LSRR options as a channel to convey              |
| the Service Chaining Header is NOT RECOMMENDED.                      |
+----------------------------------------------------------------------+
```

[6.5](). Define a new IPv4 Option and IPv6 Extension Header

Another candidate solution to convey the Service Chaining Header is
to define a new IPv4 option [[RFC0791]()] and a new IPv6 extension header
[[RFC6564]()]. The IPv4 option/IPv6 extension header can be inserted by
the classifier following the pre-configured classification rules.
Note:

o   This proposal is valid for any transport protocol.

o   This proposal offers the same functionality in both IPv4 and IPv6.

o   Some general recommendations documented at
    [I-D.ietf-opsec-ip-options-filtering], [RFC6192], and
    [I-D.ietf-6man-ext-transmit] are to be taken into account.
    Nevertheless, these security threats do not apply for this usage
    since the Ingress Node is the entity that is responsible for
    injecting the new option. Therefore, malicious usage of this
    option is unlikely.

o   Injecting an additional IP option by the classifier introduces
    some implementation complexity in the following cases: The packet
    is at or close to the MTU size, and the option space is exhausted.

o   The option can be designed to be compact and therefore avoid
    inducing fragmentation.

o   Despite it is widely known that routers and middleboxes filter IP
    options (e.g., drop IP packets with unknown IP options, strip
    unknown IP options, etc.), this concern does not apply for the
    Service Function Chaining case because the support of new IP
    options can be enabled within a domain operated by the same
    administrative domain.

o   Intermediary Nodes must not strip this IPv4 option/IPv6 extension
    header.

o   The use of an IPv4 option or IPv6 Extension Header to drive the
    processing of an incoming packet may alter the performance of SF
    Nodes.

    *   Some vendors claim the use of Extension Headers (other than
        Hop-by-Hop) does not impact the overall performance of their
        IPv6 implementation (e.g., [Report]).
    *   Some studies revealed an increase of the single-hop delay when
        IP options are included (e.g., [Delay]).
    *   The severity of the overall performance degradation is to be
        further assessed ([RFC5180]).

o   Carrying the Service Chaining Header as an IPv4 option/IPv6
    extension header allows to:

    *   De-correlate packet marking from forwarding constraints.
    *   Avoid requiring an internal tagging mechanism to each SF Node
        to preserve the same marking in the outgoing interface as the
        one received through the incoming interface.

```
+----------------------------------------------------------------------+
| Proposed Recommendation                                              |
+----------------------------------------------------------------------+
| Define a new IPv4 option and IPv6 extension header                   |
| as an Experimental track RFC document. This approach is pragmatic,|
| assuming further experiments can be conducted to:                    |
|                                                                      |
```

```
| 1. Assess the impact on performance.                            |
|                                                                 |
| 2. Compare the impact of using the IPv4 option and the IPv6     |
|    extension header vs. an encapsulation mode (i.e., in contexts|
|    where no encapsulation is required to reach the next SF hop). |
|                                                                 |
| 3. Assess to what extent the use of an IPv4 option/IPv6 extension|
|    header simplify internal tagging mechanisms specific to each SF|
+-----------------------------------------------------------------+
```

6.6. Define a New TCP Option

This proposal consists in defining a new TCP option to convey the
Service Chaining Header. The drawbacks of this proposal are listed
below:

o Encapsulating every received packet in TCP SYN messages may impact
   the performance of SF nodes.

o Injecting a TCP option by intermediate nodes will interfere with
   end-to-end (E2E) issues. One example of such interference would
   be terminating and re-originating TCP connections not belonging to
   the transit device.

o Injecting this TCP option introduces some implementation
   complexity if the options space is exhausted. TCP option space is
   limited and might be consumed by the TCP client.

o SF Nodes may need to maintain a lot of state entries to handle
   fragments.

```
+----------------------------------------------------------------------+
| Proposed Recommendation                                              |
+----------------------------------------------------------------------+
| Defining a new TCP option as a channel to convey the Service         |
| Chaining Header is NOT RECOMMENDED.                                   |
+----------------------------------------------------------------------+
```

6.7. Use the GRE Key

[RFC2890] defines key and security extensions to GRE (Generic Routing
Encapsulation, [RFC2784]). GRE Key and sequence number fields are
optional. This section investigates how a GRE Key optional field can
be used to convey a 32-bit SF Map Index.

o GRE Checksum and Sequence Number fields are not required. These
   fields must not be included.
o Relying on GRE optional field to drive the processing of received
   packets may impact the performance of SF Nodes.
o This proposal does not allow to convey additional information than
   the SF Map Index (if needed).
o In cases where GRE would already have been used, it is preferable
   to rely on this scheme and avoid yet another encapsulation
   overhead.
o An SF Node must rely on an internal tagging procedure to preserve
   the same header be positioned at the outgoing interface of an SF
   node.
o Further experiments may be required to compare the performance
   that would result in activating this solution vs. the performance
   observed when an IPv4 option or IPv6 extension header is used
   jointly with IP-in-IP encapsulation [RFC2003].

```
+--------------------------------------------------------------------+
| Proposed Recommendation                                            |
+--------------------------------------------------------------------+
| To be completed                                                    |
+--------------------------------------------------------------------+
```

6.8. Define a New IP-in-IP Scheme

This proposal is compliant with [RFC1853]. It consists in adding a
fixed header as shown in Figure 5:

```
                                   +---------------------------+
                                   |      Outer IP Header      |
                                   +---------------------------+
                                   |        SFC Header         |
     +---------------------------+ +---------------------------+
     |        IP Header          | |      Inner IP Header      |
     +---------------------------+ ====> +---------------------------+
     |                           | |                           |
     |        IP Payload         | |        IP Payload         |
     |                           | |                           |
     +---------------------------+ +---------------------------+
```
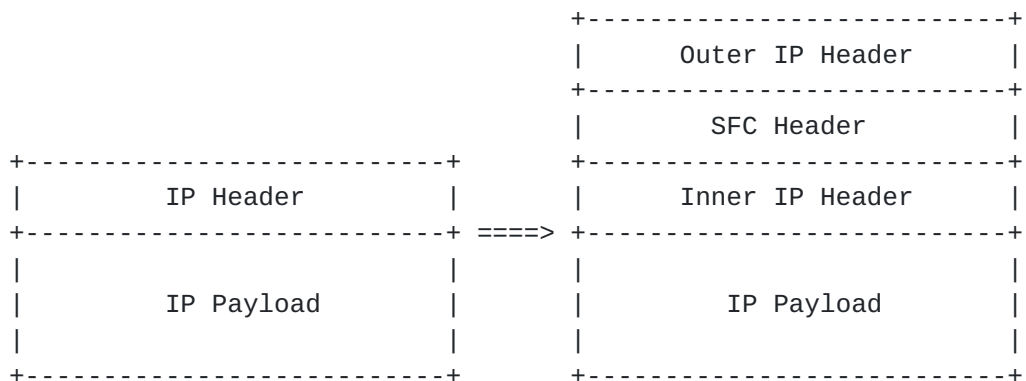
Figure 5

The following comments can be made:

o This proposal covers both IPv4 and IPv6 deployment cases.

o An SF Node must rely on an internal tagging procedure to preserve
   the same header be positioned at the outgoing interface of an SF
   node.

o This header can be extended easily to accommodate new
   requirements.

o Because the SFC Header is part of the mandatory header, the
   performance are likely to not be severely impacted compared to
   other tunneling modes such as the joint use of IP-in-IP and an
   IPv4 option/IPv6 extension header.

```
+---------------------------------------------------------------------+
| Proposed Recommendation                                             |
+---------------------------------------------------------------------+
| To be completed                                                     |
+---------------------------------------------------------------------+
```

6.9. MAC-based SFC Forwarding

The SFC Classifier capability introduced in
[I-D.boucadair-sfc-framework] can be for instance supported by a GGSN
node of a mobile network that also embeds the PCEF (Policy Charging
Enforcement Function) function. A generic description of the related
Gi Interface use case is discussed in [I-D.liu-sfc-use-cases].

This candidate solution assumes the following:

o The SFC Classifier node is connected to various SF Nodes via a
   tunnel (e.g., VxLAN or L2VPN tunnel).

o A large block of MAC addresses are allocated to the SFC Classifier
   node.

o The SFC Classifier node can use different MAC addresses in the
   Source Address field of the data frame to identify different SFCs.

o Out-of-band message(s) can be exchanged between a SFC Classifier
   node and SF Nodes to signal the SFC associated to each Source MAC
   address.

This proposal is a particular case of the "Router-based mechanism"
defined in Section 10.2 of [I-D.boucadair-sfc-framework].

The following comments can be made for this candidate solution:

o It can apply to any transport protocol.

o A large block of MAC addresses has to be allocated to the SFC
   Classifier node. The SFC Classifier node must have the extra
   logic of using different MAC addresses for different SF chains.

o Each SF node needs to be provisioned with instructions or policies
   provided to and relayed by the classifier on where to send a
   packet based on the source MAC address associated to a specific
   SFC.

o It is designed for topologies where Classifier and SF nodes can be
   connected by tunnels to maintain their Layer 2 connections. In
   particular, these tunnels are used to convey SFC-specific
   instructions and policies to the SF nodes. From that standpoint,
   the proposal is only applicable to such topologies.

o The proposed scheme requires that all traffic traverses the
   Classifier node first. The return path doesn't have to go back to
   the SFC Classifier node because all the SF Nodes forward traffic
   based on the instructions and policies provided to and relayed by
   the Classifier, instead of making forwarding decisions based upon
   the destination addresses.

o When multiple SFs that are part of a given SFC are co-located in
   the same device, the SFC Classifier may have trouble to decide
   which SF needs to be invoked in which order. A solution to avoid
   such complication is to use different source addresses to indicate
   which SFs and in which order. SF nodes (or Proxy Nodes) may need
   policies from the PDP, classification nodes, or control plane on
   how to steer packets based on source address to their designated
   SFs.

o The mechanism may be exposed to an overlapping MAC address
   situation whenever some of these MAC addresses need to be locally
   administered.

```
+--------------------------------------------------------------------+
| Proposed Recommendation                                            |
+--------------------------------------------------------------------+
| The MAC-based SFC forwarding is designed for specific topologies   |
| and assumes strong requirements on the SFC Classifier Node.        |
|                                                                    |
| The use of MAC-based SFC forwarding is NOT RECOMMENDED as a        |
| generic SFC mechanism that would remove dependency on the          |
| underlying physical topology.                                      |
+--------------------------------------------------------------------+
```

7. Steer Paths To Cross Specific SF Nodes

7.1. Need for a Mandatory Encapsulation Scheme

For interoperability reasons, one encapsulation mode MUST be defined.
Refer to [RFC3439] for more discussion on the design principles.

7.2. Candidate Solutions

Given the requirements identified in
[I-D.boucadair-sfc-requirements], IP-based encapsulation schemes
should be considered. From this standpoint, the following
encapsulation candidate solutions are identified so far:

1. Simple IP-in-IP & a SFC header in the inner packet (e.g., IPv4
    option, IPv6 extension header)
2. IP-in-IP with a fixed SFC header (Section 6.8).
3. GRE & GRE Key as a channel to convey the SF Map Index
    (Section 6.7)

7.3. Discussion

The following table summarizes the main characteristics for each
mode:

Mode	Simple IP-in-IP & a SFC header in the inner packet	IP-in-IP with a fixed SFC header	GRE & GRE Key
Encapsulation overhead when the next hop SF is in the same subnet	No	Yes	Yes
A proprietary internal tagging mechanism is required	No	Yes	Yes
Natural extensibility	Yes	Yes	No
Risk to strip the header by intermediate nodes	Yes	No	No
Possible Impact on Performance	Med to High	Low to Med	Med

The following comments can be made:

o Both "IP-in-IP with a fixed SFC header" and "GRE & GRE Key"
   present almost the same characteristics except "IP-in-IP with a
   fixed SFC header" can be easily extended. Note, "GRE & GRE Key"

can also be extended with new optional fields but this may induce
some performance degradation.
o "Simple IP-in-IP & a SFC header in the inner packet" is more
flexible:

   * It allows to convey the SFC header separately from the
      encapsulation header.
   * It allows to avoid encapsulation overhead when adjacent SFs in
      a SFC sequence are in the same subnet.
   * No internal tagging is needed within a SF Node.
   * The SFC header can be extended in the future (if needed).
o Indicated values for "Possible Impact on Performance" are
hypothetical. These values are inspired from some experiments
such as [Delay]. Ideally, further testing should be conducted to
better qualify the impact on performance of these proposals under
the same configuration and setup.

```
+----------------------------------------------------------------------+
| Proposed Recommendations                                             |
+----------------------------------------------------------------------+
| (1) Adopt the IP-in-IP with a fixed SFC header solution (Section     |
| 6.8). This mode is to be used as the MANDATORY encapsulation         |
| scheme for service chaining purposes. The main selection criteria    |
| for this proposed recommendation is to minimize performance          |
| impacts on involved nodes.                                           |
|                                                                      |
| (2) To accommodate deployment cases where encapsulation is not       |
| required, allow to rely exclusively on a dedicated tagging field     |
| in the inner packet. This extension is to be defined in the          |
| EXPERIMENTAL track (e.g., Section 6.5).                              |
|                                                                      |
| (3) Experimental specifications can be obsoleted or promoted to      |
| be in the Standard Tracks based on the conclusions from              |
| significant experiments.                                             |
+----------------------------------------------------------------------+
```

8. Summary

As a consequence of the above analysis, the following recommendations
are made:

o **** TO BE COMPLETED ONCE THE ANALYSIS IS STABLE ****

9. IANA Considerations

Authors of this document do not require any action from IANA.

10. Security Considerations

   Security considerations related to Service Function Chaining are
   discussed in [I-D.boucadair-sfc-framework].

11. Acknowledgements

   Thanks to J. Halpern for the discussion on the subscriber-ID.

12. References

12.1. Normative References

   [RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September
              1981.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax
              Used to Form Encoding Rules in Various Routing Protocol
              Specifications", RFC 5511, April 2009.

12.2. Informative References

   [Delay]   Papagiannaki, K., Moon, S., Fraleigh, C., Thiran, P., and
              C. Diot, "Measurement and Analysis of Single-Hop Delay on
              an IP Backbone Network", August 2003.

   [I-D.boucadair-intarea-host-identifier-scenarios]
              Boucadair, M., Binet, D., Durel, S., Chatras, B., Reddy,
              T., and B. Williams, "Host Identification: Use Cases",
              draft-boucadair-intarea-host-identifier-scenarios-03 (work
              in progress), March 2013.

   [I-D.boucadair-sfc-framework]
              Boucadair, M., Jacquenet, C., Parker, R., Lopez, D.,
              Guichard, J., and C. Pignataro, "Service Function
              Chaining: Framework & Architecture", draft-boucadair-sfc-
              framework-00 (work in progress), October 2013.

   [I-D.boucadair-sfc-requirements]
              Boucadair, M., Jacquenet, C., Jiang, Y., Parker, R., and
              C. Pignataro, "Requirements for Service Function
              Chaining", draft-boucadair-sfc-requirements-00 (work in
              progress), October 2013.

   [I-D.ietf-6man-ext-transmit]

          Jiang, S., "Transmission and Processing of IPv6 Extension
          Headers", draft-ietf-6man-ext-transmit-05 (work in
          progress), October 2013.

   [I-D.ietf-opsec-ip-options-filtering]
          Gont, F., Atkinson, R., and C. Pignataro, "Recommendations
          on filtering of IPv4 packets containing IPv4 options.",
          draft-ietf-opsec-ip-options-filtering-05 (work in
          progress), September 2013.

   [I-D.liu-sfc-use-cases]
          Liu, W., Li, H., Huang, O., Boucadair, M., Leymann, N.,
          Cao, Z., and J. Hu, "Service Function Chaining Use Cases",
          draft-liu-sfc-use-cases-00 (work in progress), October
          2013.

   [I-D.quinn-sfc-problem-statement]
          Quinn, P., Guichard, J., Surendra, S., Agarwal, P., Manur,
          R., Chauhan, A., Leymann, N., Boucadair, M., Jacquenet,
          C., Smith, M., Yadav, N., Elzur, U., Nadeau, T., Gray, K.,
          McConnell, B., and K. Kevin, "Service Function Chaining
          Problem Statement", draft-quinn-sfc-problem-statement-01
          (work in progress), October 2013.

   [RFC1853] Simpson, W., "IP in IP Tunneling", RFC 1853, October 1995.

   [RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003,
          October 1996.

   [RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black,
          "Definition of the Differentiated Services Field (DS
          Field) in the IPv4 and IPv6 Headers", RFC 2474, December
          1998.

   [RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z.,
          and W. Weiss, "An Architecture for Differentiated
          Services", RFC 2475, December 1998.

   [RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P.
          Traina, "Generic Routing Encapsulation (GRE)", RFC 2784,
          March 2000.

   [RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE",
          RFC 2890, September 2000.

   [RFC3439] Bush, R. and D. Meyer, "Some Internet Architectural
          Guidelines and Philosophy", RFC 3439, December 2002.

   [RFC5180] Popoviciu, C., Hamza, A., Van de Velde, G., and D.
              Dugatkin, "IPv6 Benchmarking Methodology for Network
              Interconnect Devices", RFC 5180, May 2008.

   [RFC6192] Dugal, D., Pignataro, C., and R. Dunn, "Protecting the
              Router Control Plane", RFC 6192, March 2011.

   [RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme,
              "IPv6 Flow Label Specification", RFC 6437, November 2011.

   [RFC6438] Carpenter, B. and S. Amante, "Using the IPv6 Flow Label
              for Equal Cost Multipath Routing and Link Aggregation in
              Tunnels", RFC 6438, November 2011.

   [RFC6564] Krishnan, S., Woodyatt, J., Kline, E., Hoagland, J., and
              M. Bhatia, "A Uniform Format for IPv6 Extension Headers",
              RFC 6564, April 2012.

   [RFC6674] Brockners, F., Gundavelli, S., Speicher, S., and D. Ward,
              "Gateway-Initiated Dual-Stack Lite Deployment", RFC 6674,
              July 2012.

   [RFC6864] Touch, J., "Updated Specification of the IPv4 ID Field",
              RFC 6864, February 2013.

   [RFC6967] Boucadair, M., Touch, J., Levis, P., and R. Penno,
              "Analysis of Potential Solutions for Revealing a Host
              Identifier (HOST_ID) in Shared Address Deployments", RFC
              6967, June 2013.

   [Report]   Cisco, "IPv6 Extension Headers Review and Considerations",
              , <http://www.cisco.com/en/US/technologies/tk648/tk872/
              technologies_white_paper0900aecd8054d37d.pdf>.

   [Robust]   Walter Willinger, W. and J. Doyle, "Robustness and the
              Internet: Design and evolution", March 2002, <http://
              netlab.caltech.edu/publications/
              JDoylepart1_vers42002.pdf>.

Authors' Addresses

   Mohamed Boucadair
   France Telecom
   Rennes 35000
   France

   EMail: mohamed.boucadair@orange.com

Christian Jacquenet
France Telecom
Rennes   35000
France

EMail: christian.jacquenet@orange.com

Ron Parker
Affirmed Networks
Acton,   MA
USA

EMail: Ron_Parker@affirmednetworks.com

Linda Dunbar
Huawei Technologies
5430 Legacy Drive, Suite #175
Plano   TX
USA

EMail: linda.dunbar@huawei.com