Network Working Group   P. Bottorff
Internet Draft   D. Fedyk
Intended status: Informational   HP Enterprise
                                                        H. Assarpour
                                                            Broadcom
Expires: July 2016   January 20, 2016

Ethernet MAC Chaining
draft-fedyk-sfc-mac-chain-01.txt

Abstract

   This document introduces and describes a simple and highly scalable
   service function chaining mechanism called MAC chaining which is
   built largely on existing Ethernet frame and forwarding capabilities.
   MAC chaining uses IEEE 802 Media Access Control (MAC) addresses to
   provide flexible and complete service function chains. It is largely
   transparent to layers above Ethernet and designed to augment and
   coexist with existing virtual and physical network forwarding. MAC
   chaining is achievable in some devices and virtual switches today
   using existing protocols.

Status of this Memo

Table of Contents

[1]. Introduction

Service Function Chaining (SFC) enables the creation of composite
(network) services that consist of a directed graph of Service
Functions (SF) which must be applied to packets selected as a result
of classification. SFC is described in detail in the SFC architecture
document [RFC7665], and is not repeated here.

This document describes a new highly scalable, low resource, service
function chain (SFC) mechanism called MAC chaining that is based on
the current IEEE 802 [802-2001] Ethernet header for physical and
virtualized environments. Service function chaining is an active area
in the standards and various proposals for how to do SFCs are being
put forward. The basic mechanism used for MAC chaining is the use of
MAC addresses in the Ethernet header as a mechanism both for
identifying chains and for forwarding packets along a MAC chain. The
forwarding mechanism used in MAC chaining is independent from virtual
or overlay networks used to form subnets. MAC chaining addresses are
terminated at each Service Function Forwarder (SFF) and replaced by a
new set of MAC chaining addresses used to forward through the next
Service Function in the chain.

```
            +-----------------------------------------+
           /                 E2E Network             /
          /                                         /
       +-----------------------------------------+
             +----------------------------------+
 O     /|   /      MAC Function Chaining       /
 R    / |  /                                  /
 C   /  | +--------------------------------+
 H  /   |    +----------------------------------+
 E |    |   /        Virtual Networks          /
 S |    |  /         Overlay/Underlay         /
 T |    | +--------------------------------+
 R |    |    +--------------+  +--------------+
 A |    |   /   Physical   / /    Physical   /
 T |    |  /   Networks   / /    Networks   /
 I |    | +--------------+  +--------------+
 O |   /
 N |  /
   | /
   |/
```

              Figure 1 Service Forwarding Plane

MAC chaining can be viewed as a network service plane as shown in
Figure 1. The SFC architecture document [RFC7665] describes chain
forwarding in terms of 3 main architecture components which are the
Service Classification Function (SCF), Service Function Forwarder
(SFF) and the Service Function (SF). When managed with MAC chaining,
Service Functions (SF) are simple links in the service chain and
require little context of the overall chain. MAC chaining Service
Function Forwarders (SFF) enable the chain and control the path to
and from the SFs. Logically the SFF forms a switching layer above the
existing virtual networking layers. In MAC chaining, a Chain
Termination Function (CTF) is added to the architecture to separate
the operation of de-encapsulating the packet and sending it toward
the final destination from the operations of service function
classification and service function forwarding described in the I-
I.ietf-sfc-architecture.

MAC Chain forwarding is performed by a MAC Chaining Service Function
Forwarder (SFF) using DA and SA address swapping. The operation of a
MAC Chaining SFF has characteristics of a router in that it uses
information in the packet to determine a new link destination,
however unlike a router the new link decision is based on the
previous MAC address rather than the IP address. This arrangement has
the advantage that the IP addresses retains the end-to-end address
eliminating the need for NAT addresses on entry and exit of the
chain. A MAC Chain Service Function Forwarder also has
characteristics of a Bridge in that it uses a promiscuous receiver
with exact matching of every frame presented on it's links to a MAC
DA and VLAN entry in the filtering database. This matching prevents
forwarding frames which don't contain allocated Chain MAC addresses.
The exact matching performed by a MAC Chaining SFF provides orders of
magnitude higher table scaling than longest match forwarding
characteristic of routers.

2. Conventions used in this document

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].

In this document, these words will appear with that interpretation
only when in ALL CAPS. Lower case uses of these words are not to be
interpreted as carrying RFC-2119 significance.

3. Terminology

Chain Termination Function (CTF): See [RFC7665]. The chain
                termination function terminates a Service Function Path

performing any de-encapsulation and operations required
to continue forwarding to the final destination. The
CTF may also be the final destination of the chain.

CS-MAC: A MAC address which identifies a MAC Chain Segment

CS-MAC Authority: CS-MAC Authority refers to the purely
administrative mechanism to ensure CS-MACs are unique
but allows the optional reuse of MACs in different VNs.
Each VN port has a single CS-MAC Authority. Multiple
ports may share the same Authority. A MAC Chain may be
under a single CS-MAC Authority or it may be split
among multiple CS-MAC Authorities.

DA: MAC destination address

I/G   The Individual/Group (I/G) address bit (LSB of octet
0).

MAC Address: IEEE 802 Media Access Control Address a 48 bit address.

MAC Chain Segment (CS): A hop between either Service Forwarding
Functions, a Service Classification Function and a
Service Forwarding Function, or a Service Forwarding
Function and a Chain Termination Function

MTU: Maximum Transmission Unit. Layer 2 has a maximum frame
size and Layer 3 has a Maximum Packet Unit. This
documents uses the term Layer 2 MTU to identify that
MAC chaining does not affect L3 or IP MTU.

VN Port: In this document a port is the logical interface
context for a MAC address in a virtual network (VN). A
VN port may be implemented on any type of physical port
or logical supporting Ethernet.

SA: MAC source address

Service Function (SF): See [RFC7665].

Service Classification Function (SCF):
See [RFC7665].

Service Function Chain (SFC): See [RFC7665].

Service Function Forwarder (SFF): See [RFC7665].

Service Function Path (SFP): See [RFC7665].

U/L: The Universally or Locally administered (U/L) address
                bit is the bit of octet 0 adjacent to the I/G bit.

Virtual Network (VN):   A Virtual network is used to identify a
                network segment controlled by a CS-MAC Authority.

4. MAC Chaining

MAC chaining uses controlled assignment of Ethernet 48 bit MAC
addresses to form the chain. Ethernet MAC addresses are selected to
uniquely identify both the chain and the particular chain segment (or
hop) within the identified chain. These assigned Ethernet addresses
are called Chain Segment MAC (CS-MAC) addresses in this document.
These CS-MACs allow MAC chaining to be implemented on existing
Ethernet infrastructure making it broadly interoperable with the
majority of installed base including existing Ethernet, Carrier
Ethernet and IP equipment.

Each MAC chain is composed of a series of Chain Segments (CS) which
are hops between Classifiers, Service Function Forwarders and Chain
Terminating Functions (see figure 2). Some of the chain segments
include Service Functions while others perform forwarding between the
SCF, SFF and CTF. For each chain segment, a Destination MAC address
(DA), and optionally a Source MAC address (SA) are selected, from a
locally administered MAC address space, to uniquely identify the
chain segment within the SFC domain. MAC chaining uses these CS-MACs,
in the Ethernet header, as an identifier to enable forwarding packets
in a MAC chain.

```
+-----+  +-----+ +------+ +-----+ +-----+ +------+ +-----| +-----+
|     | |      | \ SF2/ |     |  | |      | \ SF4/ |     | |     |
| SCF +--+ SFF1+---+B +--+ SFF1+-+ SFF2+---+D +---+ SFF2+-+ CTF |
|    X| 1|A   C|2   \/   2|C    | |    E|3   \/   4|E   Y| |F    |
+-----+  +-----+          +-----+ +-----+          +-----+ +-----+
   \...../   \.........../   \.../   \.........../   \..../
     CS 1         CS 2         CS 5         CS 3         CS 4
   CS-MAC=A      CS-MAC=C    CS-MAC=G    CS-MAC=E     CS-MAC=F
```
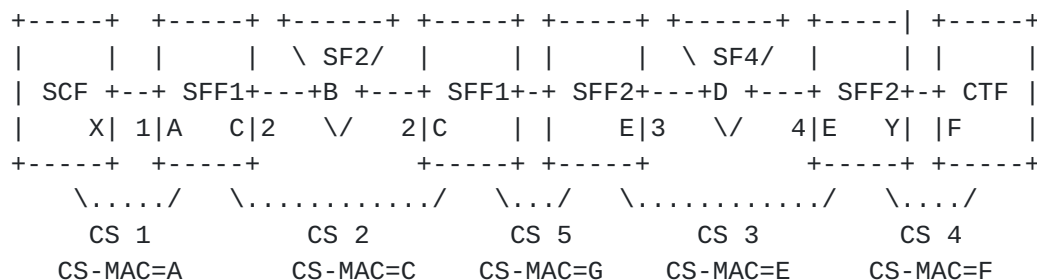
Figure 2 MAC Chain Segments Addressing

In Figure 2 five chain segments are illustrated. The first chain
segment is between the classifier and service function forwarder
identified as SFF1. This chain segment, designated CS1, has been

assigned CS-MAC A. (For brevity the 48 bit MAC addresses are
identified by letters). The next chain segment is from SFF1 VN port 2
through service function 2 and back to SFF1 VN port 2. This chain
segment designated CS2 has been assigned CS-MAC C. SF2 on CS2 is a
single armed SF with MAC address B attached to SFF VN port 2. (See
section 4.3.3 for a description of the types of armed SF). Chain
segment CS5 is between SFF1 and SFF2. It has an assigned CS-MAC G.
Chain segment CS3 from SFF2 VN port 3 to SFF2 VN port 4 is identified
by CS-MAC E. SF4, lying on CS3, is a dual armed SF with MAC address D
on the side connecting to SFF2 VN port 3. The final chain segment,
CS4, of the path is between SFF2 and the CTF and is identified by CS-
MAC F.

As described here, MAC chaining operates in the context of Virtual
Networks (VN). To fully describe each MAC chaining address the tuple
(CS-MAC Authority, CS-MAC) is used which uniquely identifies each
chain segment as well as the entire chain. Each chain segment and
each VN MUST belong to a single CS-MAC Authority which is a
management construct that assigns unique CS-MACs for that segment and
VN. If a chain segment crosses between two independent VNs, then
both the VNs must have the same CS-MAC Authority.

4.1. MAC Chaining Packet and Address Formats

The IEEE 802.3 [802-2001] frame header consists of a Destination MAC
Address (DA)(6 bytes) followed by a Source MAC Address (SA)(6 Bytes)
followed by a number of possible fields which are identified by an
Ethertype (2 Bytes) following the SA. A VLAN tag (4 bytes) is a
common TAG that also carries Priority Code Points and Discard
Eligible Information for traffic classification.   For the purpose of
this document the DA and SA are the primary fields used for MAC
chaining however the frame may optionally have VLAN Tags. The MAC
chaining frame can also be carried inside other encapsulations (i.e.
within an overlay) like VxLAN, L2VPN or Provider Backbone Bridging
(PBB).

Figure 3 illustrates the formats MAC chaining uses to carry the
original IPv4 or L2 packets when entering the classifier. Since MAC
chaining encodes a SFC path solely in the MAC addresses of the
Ethernet header the SPI and SI fields of the Network Service Header
(NSH) [I-D.ietf-quinn-sfc-nsh] are not necessary and therefore NSH is
an optional addition when using a MAC segment chain.

```
Format 1: Original IPv4, MAC Chaining without NSH:
+--------------------------------------+-------------------+
| Outer Ethernet, ET=0x0800            | original IP Packet |
+--------------------------------------+-------------------+


Format 2: Original IPv4, MAC Chaining with NSH
+--------------------------+--------------+-------------------+
| Outer Ethernet, ET=0x894F | NSH, NP = 0x1 | original IP Packet |
+--------------------------+--------------+-------------------+


Format 3: Original L2, MAC Chaining without NSH
+--------------------------------------+-------------------+
| Outer Ethernet, ET=0x****            | original L2 frame  |
+--------------------------------------+-------------------+


Format 4: Original L2, MAC Chaining with NSH
+--------------------------+--------------+-------------------+
| Outer Ethernet, ET=0x894F | NSH, NP = 0x3 | original L2 frame  |
+--------------------------+--------------+-------------------+
```
      Figure 3 MAC Chaining Formats for IPv4 and L2 Service Packets

For original L3 packets MAC Chaining can forward a standard L3 frame
without any further encapsulation. In addition if the SFs or Proxy
functions are NSH aware, the format 2 for original L3 packets allows
adding the NSH header to pass meta-data between SFs.

Figure 3 provides two alternate encapsulations for original L2
packets. The simple format 3 encoding without NSH uses an Ethertype
to designate an L2 frame follows. This encoding provides an L2
encapsulated in an L2 frame. The format 4 encoding uses the NSH
header with the Next Protocol set to 0x3 designating an L2 frame
encapsulation. The NSH encapsulation provides all the functions of
the simple L2 encapsulation and therefore can be used whenever the
SFs or Proxy functions are NSH aware.

In addition to meta-data carried in the NSH it is also possible to
encode a small amount of meta-data in the Chain Segment MAC
addresses. The branch taken bit (figure 4) is a small piece of meta-
data that can be used in the MAC chaining header. For more elaborate
meta-data, the Network Service Header draft [I-D.ietf-quinn-sfc-nsh]
header is compatible with MAC chaining. Section 11.3 of [I-D.ietf-
quinn-sfc-nsh] illustrates that the Ether type following a MAC
chaining outer header has a registered type of 0x894F(TBC) with an
NSH that subsequently defines the payload. SFs can use the NSH within
the chain. Any SCF, SF or CTF can remove or modify the NSH as

specified in the NSH draft. When using the IETF NSH draft each SF
must either be capable of receiving an Ethernet frame with the NSH or
must be supported by a proxy which removes the NSH before the SF.

The format of the MAC address used by MAC chaining is the standard
IEEE MAC address format of 48 bits as illustrated in figure 4.

Every MAC address is identified as either a global or a local MAC
address. Global MAC addresses are intended to be worldwide unique
while local address are intended for the use of local administrations
domains and are not worldwide unique. Each global address uses a 22
bit Organizational Unique Identifier (OUI) prefix which is assigned
by the IEEE Registration Authority Committee to support worldwide
uniqueness. Recently, in response the needs of virtualization
environments, the IEEE Registration Authority Committee has started
assigning 22 bit Company IDs (CID) to allow independent vendors to
share the local MAC addresses space within a domain where multiple
un-coordinated authorities are assigning addresses. MAC chaining MAY
make use of the new administered local MAC addresses.

```
 I U
 / /                                                           B
 G L                                                           T
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |X|n|CID1[42:47]|  CID2[32:39]  |  CID3[24:31]  |CS-ID[16:22] |X|
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | CS-ID [8:15]  |  CS-ID [0:7]  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
        Figure 4 MAC Chaining Ethernet MAC address Format

Where:

I/G   IEEE 802 Ethernet Individual/ Group address bit.

U/L   IEEE 802 Ethernet Universal / Local Bit. Bit MAY be set
         indicating local.

CID   Company ID - 22 Bits (Not Mandatory example only). Company
         Ids are assigned by IEEE registration to vendors who use
         Local addresses for MAC chaining or other purposes. The CIDs
         are unique and ensure that there are no collisions with
         other protocols that use local addresses. However the local
         addresses can be reused in other networks.

BT   Branch Taken Chain indicator. Required. This bit may be set
        or reset in context of a chain.

CS-ID   Chain Segment ID 23 bits. (Example only).

CS-MAC addresses, which identify chain segments, SHOULD be allocated
by the CS-MAC Authority from the local space using a Company ID
assigned to the CS-MAC Authority. MAC addresses also have an
Individual (unicast) or Group (Multicast) bit I/G. MAC chaining MAY
use individual or group addresses for the CS-MACs though restriction
on the use of group CS-MACs may apply depending on the type of
forwarding performed by the SFF for the particular segment.

MAC chaining may also use global or local MAC addresses. The MAC
address assigned to a Service Function MAY be Global or Local and can
be assigned by any authority, not necessarily the CS-MAC Authority.

As with other types of Service chaining, a packet or a frame travels
through a network until it encounters an initial classifier.
Forwarding before the classifier is out of the scope of this
specification. The native packet format (L2 or L3 or tunneled, etc.)
arriving at the classifier does not matter but the classifier (or set
of classifiers) need to inspect the packet and determine that the
packet is part of a service chain.

In all cases of MAC chaining after a frame (L2, L3, etc) has been
classified the MAC chain begins by prepending the packet with an
Ethernet L2 Frame header. The frame will also have a valid 4 byte CRC
checksum.

One advantage of MAC chaining is the MAC frame has an overhead of
bytes that can leave the L2 MTU unaffected. As with all Ethernet II
frames payload must be a minimum of 64 bytes or must be padded to 64
bytes.

4.2. Meta-Data Encoding Consideration for the Network Service Header

A NSH is required on NSH aware chain segments where meta-data is
being carried. A NSH may be inserted and deleted from the chain
depending on the requirements of the specific SFs by the MAC Chaining
SFFs as discussed in 4.1.

The current NSH draft [I-D.ietf-quinn-sfc-nsh] has a mandatory 32
byte service path header. Since the MAC Chaining SFFs don't use the
two fields of this service path header: Service path and Service
index, of the NSH for forwarding, when NSH is included these fields
may be used for other meta-data. Meta-data is likely to have many

forms based on specific chain segments so an additional 32 bits of
flexible meta-data may prove useful. Figure 4 shows an alternative
NSH header format that is field compatible with the existing draft
fixed header but allows additional flexibility when used with SFC
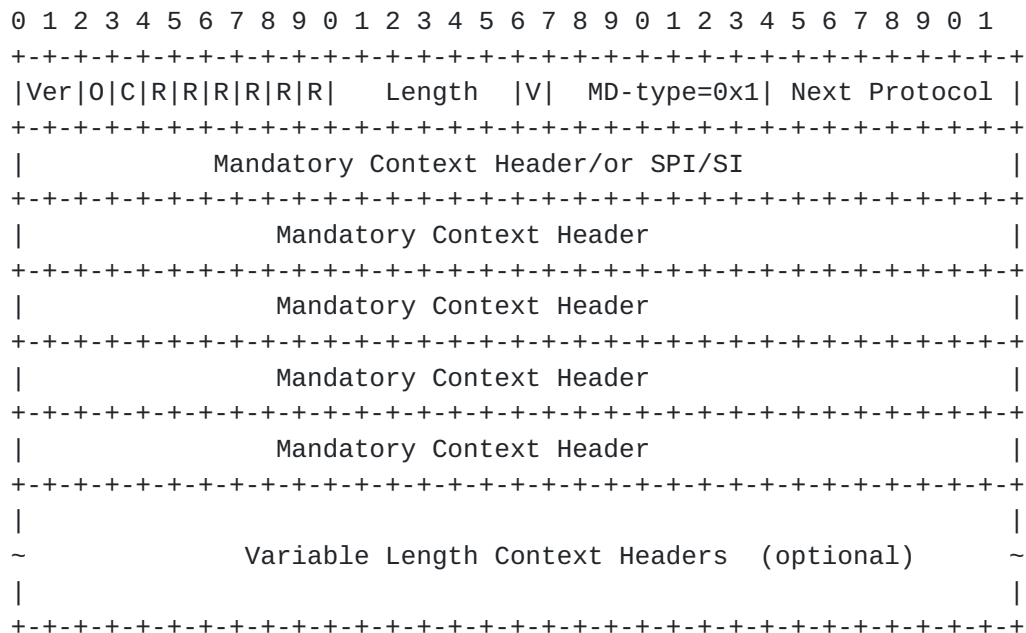transport that do not require facilities.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Ver|O|C|R|R|R|R|R|R|   Length  |V|  MD-type=0x1| Next Protocol |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Mandatory Context Header/or SPI/SI                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Mandatory Context Header                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Mandatory Context Header                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Mandatory Context Header                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Mandatory Context Header                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              |
~              Variable Length Context Headers  (optional)     ~
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: Proposed modifications to MAC Chaining NSH

For the Base Header Field Descriptions see [I-D.ietf-quinn-sfc-nsh].

If additional Meta-data is required beyond the fixed header we
propose the variable length structure can simply be concatenated as a
consistent method for adding TLV based meta-data in a way which
augments any fixed header allocations. Rather than defining MD-type 2
as variable headers only, this documents proposes variable headers
should be used when the fixed header space does not allow sufficient
space. The variable and fixed fields can be used in combination by
looking for the variable TLVs after the fixed header. This can be
determined without consulting MD-type by simply checking the length.

There are several drafts that are proposing context based headers
[I-D.meng-sfc-nsh-broadband-allocation], [I-D.guichard-sfc-nsh-dc-
allocation], and [I-D.napper-sfc-nsh-mobility-allocation]. If a MAC
chaining SFC is used there is up to an optional 32 bits of other
context that may be used before resorting to a TLV based header. In a
hybrid SFC chain environment that supports different SFC forwarding

on different chain segments, SFs could map MAC chainIDs to Service
paths if required.

4.3. Forwarding

Forwarding of a packet is from a classifier (SCF) to the Service
Function Forwarder (SFF) to the service function (SF) to the next SFF
to the next SF and so on until the chain is finished. MAC chaining
makes the distinction that the forwarding operations performed by a
SFF and a SF are distinct and independent. However implementations
may place SFF and SF functions as combined or separate entities. This
makes MAC chaining particularly useful for deployment in
virtualization environments where a virtual machine may implement one
or more SFs and SFFs. Forwarding is a table driven operation. Note
that all active chains are normally preprogrammed.

Figure 5 illustrates the table driven forwarding operation of a MAC
chaining SFF. Every frame arriving on the ingress VN port is matched
to the MAC chaining filtering database. On arrival at the SFF the DA
always contains a CS-MAC for the chain segment just crossed. The DA
is looked up in the context of Port 1 (a VN port) and the subsequent
DA prime (DA' a CS-MAC used as the new frame DA if this segment is DA
forwarding) and SA' (a CS-MAC used as the new frame SA if this
segment is DA/SA forwarding) and egress Port 1 prime (1') are
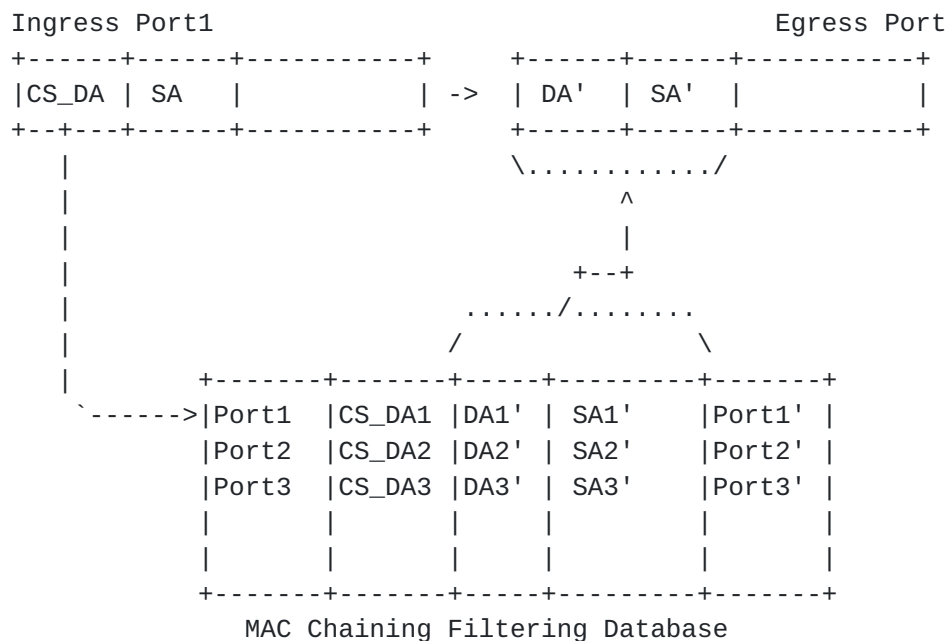determined by the lookup.

```
   Ingress Port1                                 Egress Port
   +------+------+-----------+    +------+------+-----------+
   |CS_DA | SA   |           | -> | DA'  | SA'  |           |
   +--+---+------+-----------+    +------+------+-----------+
      |                           \............/
      |                                 ^
      |                                 |
      |                               +--+
      |                      ....../........
      |                     /              \
      |        +-------+-------+-----+---------+-------+
       `------>|Port1  |CS_DA1 |DA1' | SA1'    |Port1' |
              |Port2  |CS_DA2 |DA2' | SA2'    |Port2' |
              |Port3  |CS_DA3 |DA3' | SA3'    |Port3' |
              |       |       |     |         |       |
              |       |       |     |         |       |
              +-------+-------+-----+---------+-------+
                 MAC Chaining Filtering Database
```


Figure 5 MAC Chaining Table Operation

Any frame which doesn't exactly match an entry in the MAC chaining
filtering database MUST be discarded. The filtering database itself
is configured under the control of a network controller (see section
6) which is responsible for creating the chain by programming the MAC
chaining filtering database. The MAC chaining filtering database is
an exact match database which may use existing Bridge match logic.
The exact match filtering with a hash implementation allows the
filtering database to easily scale to a large number of chains.

The Branch Taken (BT) Operation bit (figure 4) allows an SF to branch
the chain by changing the BT bit. Not all service chains are branch
capable. If a simple branch chain is desired it must be programmed in
the SFF filtering database. A branch operation is indicated by
setting a reserved bit in the CS-MAC address. This bit is not read as
a bit but as a paired address to the forward direction. The context
of the BT bit may be maintained in both the SA and the DA once the
bit has been set. An SFF receiving a frame with the BT bit set will
look up the exact match address and forward the frame in the context
of the received VN.

More complicated branching requires SF chain awareness. The next hop
addresses may be overridden by chain aware SFs to perform more
advance branching. A SF must be provided with the allocated addresses
for larger branches.

Any frame that arrives at an SFF and is not found in the forwarding
table is dropped. Devices in the path that are not MAC service
chaining aware are free to bridge the frame normally or to route
using any underlay Layer 3 or 2.5 VN encapsulation.

MAC Chaining supports two different types of forwarding methods which
are called DA forwarding and DA/SA forwarding. These two types of
forwarding are used for coupling different types of SFs into a chain.
The DA forwarding method is suited to operating on existing SFs (such
as Firewalls) which provide transparent or Bridge forwarding modes.
The DA/SA forwarding method is suited for use with Virtualized SFs
that are operating in Virtual Machines or Containers. The main
difference between the two methods is when using DA forwarding the
SFF encodes the Chain Segment MAC in the DA field. The DA therefore
contains an address for the next SFF hop rather than an explicit
address for the SF. In DA/SA forwarding the SFF encodes the Chain
Segment MAC in the SA rather than the DA. For DA/SA forwarding the
SFF uses the DA to directly address the SF. This allows any SF to
receive on a single unique address which may be shared by all chains
passing through the SF. The choice of DA or DA/SA forwarding is made
for each SF of the chain depending on the requirements of the
specific SF.

4.3.1. Destination Address MAC Chaining Operation

Destination MAC address chaining uses only the Destination MAC
address to key on and implement a chain. Destination Address MAC
chaining is used to operate with a MAC chaining un-aware SF. A
Classifier/Service Function Forwarder (SFF), composing a DA MAC
chaining hop, encodes the Chain Segment MAC in the frame DA and an
address for the SFF in the frame SA. This encoding will address the
next SFF or CTF in the chain. DA MAC chaining may only be used with
dual-arm or multi-arm SFs since an unmodified frame can't be returned
to the same network where it was received. Any SF along a DA MAC
chaining segment must be operating in Transparent or in Bridging mode
so it behaves as a "bump in the wire".

For a Service Function to participates in a DA MAC chaining it must
operate in promiscuous receiver, like an Ethernet Bridge, rather than

explicit receiver used by Ethernet stations. A MAC chaining Service
Function Forwarder uses promiscuous receiver on its' VN ports just
like every Ethernet Bridge and most Routers. In promiscuous receiver
the switch receives and inspects every frame presented to it
independent of the addressing on the frame.

DA MAC chaining is determined by the configuration of the forwarding
table in the SFF. After the initial classification the packet is
passed to the first SFF (this may be a virtual operation completely
within a single chaining switch). The SFF formats the Ethernet frame
with a DA of the next hop in the Chain. At each hop the MAC address
is looked up in a table similar to figure 2.

4.3.2. Destination and Source Address MAC Chaining

DA and SA MAC chaining is a variation of MAC chaining that allows MAC
chaining aware SFs to use explicit receiver and to support single
armed as well as dual and multi-armed SFs. When using DA/SA MAC
chaining the SF is individually addressed by the DA and therefore
does not need to operate using a promiscuous learning receiver. Such
a SF does not need a MAC lookup table and may be provisioned with a
single global or local address under any administration authority
(not necessarily the MAC chaining address authority). Service
Functions using DA/SA MAC chaining require only a single MAC address
regardless of the number of chains passing through them. DA/SA MAC
chaining is particularly advantageous for virtual service functions
(VNFs) since it reduces the need to flood frames into the virtual NIC
supporting the SFs virtual machines and server I/O accelerators.

DA/SA MAC chaining uses both addresses in the Ethernet L2 Header. The
DA is used for the next hop device and the SA is used for the
subsequent next hop device of the chain. A SF receives a frame;
processes the frame; replaces the DA with the received SA and uses
resulting DA (received SA) to forward the frame. By specifying 2 hops
in a chain the SF can be a very generic operation. The original SA of
the received frame does not have to be the address at the SFF that
created the header, allowing forwarding flexibility.

4.3.3. Forwarding by Service Functions

MAC chaining Service Functions (SFs) must be able to pass Ethernet
DA/SA addresses through the SF unless the SF is supported by a Proxy
Forwarder (see Proxy Forwarders section below). SFs are not required
to pass VLAN Tags. Service Functions supported by MAC chaining can be

classified by how they attach to the network as single armed, dual
armed or multi-armed. Single arm SFs receive and send all packets on
the same VN port. Single armed SFs are typically used when the
direction of travel is unimportant to the SF. Dual arm SFs have two
VN ports and pass packets between the two VN ports. Dual arm SFs are
typically used when the SF needs to know the direction of travel.
Multi-arm SFs have more than two VN ports. In a multi-arm (two or
more) the SF selects the egress VN port based on its' re-
classification of the packet. Each VN port of a multi-arm SF must
attach to a different VN or the SF must be MAC Chaining aware. These
SFs allow the SFs to branch the chain based on re-classification or
to replicate in the chain.

4.3.4. Proxy Forwarders

Proxy forwarding is typically for legacy devices or other devices
that do not have an ability to support MAC chaining by passing
through L2 headers.

Some service functions may reside on devices that do not understand
MAC chaining. Legacy functions on middle boxes are one example. In
this case a proxy forwarding function is used. Proxies may be
integrated with the SFF or located in the switches attaching to the
FS. The proxy will removing the MAC chaining header and forwarding
the packet in an appropriate format to the SF. The SF then returns
the packet to the proxy upon completion of its operation. The
Specific formats of frames between the proxy and the SF, when using a
proxy is out of the scope of this document.

The most basic proxy is a transparent proxy, which must be located
between the SF and any underlay entity. A transparent proxy provides
a provisioned Ethernet header which is used for forwarding all frames
egressed by a SF at a specific VN port. The use of a transparent
proxy limits the utility of the service chain in which it is inserted
since no chain state is passed through the SF by the proxy.

It is possible to proxy for chain unaware SFs by using a proxy which
stores DA/SA pairs on ingress to the SF and restoring the DA/SA on
egress from the SF. A MAC chain aware SF can determine if the packet
is using DA or DA/SA forwarding by determining if the DA the address
of the SF itself.

4.3.5. Example MAC Chaining Walk Through using DA/SA Chaining

   Figure 6 outlines the general path and operations of a MAC chaining.

   The Service Classification Function (SCF) determines if a packet
   matches a predetermined policy for the chain by inspecting the packet
   then selecting the chain by encoding the frame with next destination
   equal to the chain segment 1 MAC Address A and itself as the Source
   Address (SA) designated as H in figure 6.

   SFF1 receives a frame from the SCF with Destination Address (DA)
   equal (exact match) to A and finds the next chain segment by looking
   up A to find the next DA equal to SF2 MAC Address B and sets the SA
   equal to chain segment 2 MAC Address C.

   SF2 is a single armed Service Function which receives and sends all
   data on a single network interface. The single SF2 network interface
   normally connects to a single virtual or physical network. SF2
   receives a frame from SFF1 performs its function and then returns the
   frame to C. This process requires the SF to forward back to the
   frame's SA, by swapping DA and SA, on the same VN port.

```
                     One Arm              Two Arm
   +-----+    +-----+  +------+  +-----+  +------+  +-----| +-----+
   |     |    |     |  |  \ SF2/  |     |  |  \ SF4/  |     | |     |
   | SCF +---+ SFF1+----+B +----+ SFF1+----+D +----+ SFF2+-+ CTF |
   |    H|    |A   C|     \/     |C    |      \/x     |E   T| |F    |
   +-----+    +-----+            +-----+              +-----+ +-----+
      \...../   \.............../   \.............../   \..../
        CS 1          CS2                  CS3             CS4


      +----+    +----+   +----+    +----+   +----+   +----+
      |A,H |    |B,C |   |C,B |    |D,E |   |E,x |   |F,T |
      +----+    +----+   +----+    +----+   +----+   +----+
      Frames showing MAC DA, SA
```

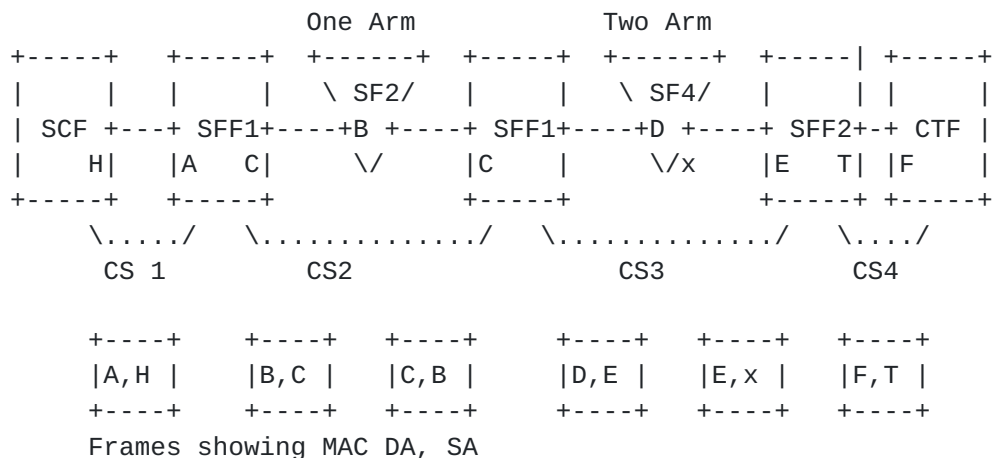   Figure 6 MAC Service Chain Example

   SCF: Service Classification Function
   CTF: Chain Termination Function
   SF: Service Function
   SFF: Service Function Forwarder
   CSx: Chain Segment x.

SFF1 receives a frame from SF2 with DA equal to chain segment 2 MAC Address C, finds the next chain segment by looking up C to find the next destination equal to SF4 MAC Address D and the SA equals chain segment 3 MAC Address E.

SF4 is a dual armed Service Function which receives and sends data on two network interfaces. SF4 always forwards frames between its two interfaces. The two interfaces of SF4 are normally connected to separate virtual or physical networks. SF4 receives the frame from SFF1 with DA equals D and SA equals E performs its function then forwards to E by swapping DA with SA and sends out the packet to the other VN port (A VN port that supports address E as a destination).

SFF2 receives a frame from SF4 with DA equals chain segment 3 MAC Address E, finds the next chain segment by looking up E to find the next destination equals chain segment 4 MAC Address F and SA equals SFF2 MAC Address T.

The CTF receives a frame from SFF2 with destination equals F. The CTF must perform any required packet header adjustment and egress VN port determination based on the destination equals F and the frame payload (i.e. uses the IP address to route the packet).

4.3.6. Forwarding by Chain Termination Functions

The forwarding to the final destination by the CTF typically does not use MAC chaining. The CTF is responsible for receiving frames addresses to the termination CS-MAC for each chain, de-encapsulating the packets, and forwarding the packets toward their final destination. One common method which may be used by the CTF for forwarding to the final destination is to route the packets using the IP address of the service packet.

If the service packet (data payload) is an L2 packet then the CTF may use either the IP network addresses or the L2 addresses to forward the packet. The choice between these two CTF forwarding models will depend on the application. Other CTF forwarding models are possible using by using the CS-MAC or meta-data for forwarding.

5. Programming a Service Chain

The capability exists today with open flow enabled switches to specify MAC match criteria and actions that match MAC forwarding all operations. However not all switches are Openflow enabled.

A Yang model could be specified to enable the MAC Chaining operations using an I2RS agent.

Chains must be preprogrammed. Care must be taken to ensure that service chain loops are not programmed (this can easily be verified before a chain is active) however MAC chains that are programmed correctly are inherently loop free in the data plane. The policy is to drop a frame that is not an exact match on any MAC chaining aware SFF.

MAC chaining may be programmed be allowed to pass through bridges that are not MAC chaining aware. It is recommended that this operation be explicitly controlled by setting up port based VLANs designed for this purpose. Ports can add a VLAN tag as part of their forwarding operation. This can be usually be achieved with existing Ethernet controls that allow ports to have service tags added. The VLAN tagging is independent of MAC chaining in this regard.

6. Considerations for Operation over NVO3 Tunnel Transports

MAC Chaining can be used with the L3 encapsulation transport tunnels being specified in NVO3 ([I-D.ietf-nvo3-vxlan-gpe], etc.). When using an NVO3 encapsulation it is preferable to use an encapsulation which supports encapsulation of an L2 packet such as [I-D.ietf-nvo3-vxlan-gpe], since this allows encoding the MAC addresses used for chain forwarding at the natural layer boundary used to address Virtual Machines or containers.

It is desirable to encode any meta-data header, such as the NSH within an NVO3 transport tunnel following an encapsulated L2 MAC header as an Ethernet tag since this will allow a natural protocol layering for delivery of the meta-data to an addressed Virtual Machine or container. Since Virtual Machines and containers are addressed by MAC addresses at the hypervisor vSwitch or system level, the meta-data will be carried as part of the frame layer into the guest OS or container environments along with the MAC addresses used for chaining.

7. Domain of operation

MAC chaining requires connectivity of L2 virtual networks over the service chain path. (This may include multiple VNs that are interconnected.) In many networks this is readily available. Data centers for example can use MAC chain within a physical site that has L2 connectivity.

If Virtualization of the L2 domain is enabled MAC chaining could
operate over L2 networks such as NVO3 or Ethernet EVPN and an
existing L2 Overlay.

8. Security Considerations

MAC chaining is an Ethernet based forwarding operation that follows
standard Ethernet rules. VN ports should be qualified with VLANs
that limit the scope of MAC chaining frames. This prevents MAC
chaining messages from being flooded to external parts of the network
or injected into a network from external sources. Programming the
VLAN that support MAC chaining is controlled and access to those
VLANs is allowed only by trusted devices.

MAC chaining is IP agnostic but like any tunneling protocol it will
deliver IP frames to other parts of a network.

9. IANA Considerations

There are no IANA considerations for this document.

10. References

10.1. Normative References

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

[802-2001] "Standard for Local and Metropolitan Area Networks:
          Overview and Architecture", IEEE 802, Standard 2014.

10.2. Informative References

[RFC7665] Halpern, J., Pignataro, C. Editors, "Service Function
          Chaining (SFC) Architecture", RFC 7665, June 2015.

[I-D.ietf-quinn-sfc-nsh]
          P.Quinn et al., "Network Service Header", draft-ietf-sfc-
          nsh-01 work in progress), July 23, 2015.

[I-D.meng-sfc-nsh-broadband-allocation]
          W. Meng et al., "NSH Context Header - Broadband", draft-
          meng-sfc-nsh-broadband-allocation-00, October 7, 2015

   [I-D.guichard-sfc-nsh-dc-allocation]
         J. Guichard et al., "Network Service Header (NSH) Context
         Header Allocation (Data Center)", draft-guichard-sfc-nsh-dc-
         allocation-03, December 14, 2015

   [I-D.napper-sfc-nsh-mobility-allocation]
           J. Napper et al., "NSH Context Header Allocation -
           Mobility", draft-napper-sfc-nsh-mobility-allocation-02,
           November 4, 2015

   [I-D.ietf-nvo3-vxlan-gpe]
           P. Quinn et al., "Generic Protocol Extension for VxLAN",
           draft-ietf-nvo3-vxlan-gpe-01. November 4, 2015

11. Acknowledgments

   This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Don Fedyk
Hewlett Packard Enterprise
153 Taylor Street
Littleton, MA
Email: don.fedyk@hpe.com

Paul Bottorff
Hewlett Packard Enterprise
8000 Foothills Blvd.
Roseville, CA
Email: paul.bottorff@hpe.com

Hamid Assarpour
Broadcom Corporation
600 Federal Street
Andover, MA
Email: hamid@broadcom.com