SFC working group
Internet Draft
Intended status: Standard Track
Expires: September 2015

L. Dunbar
A. Malis
Huawei

March 6, 2015

Framework for Service Function Path Control
draft-dunbar-sfc-path-control-01.txt

Status of this Memo

Copyright Notice

Abstract

This draft describes the framework of Service Function Path Control when some service functions on the path fail or need to be replaced.

Table of Contents

1. Introduction

This draft describes the framework of Service Function Path (SFP) control when some functions on the path fail or need to be replaced.

SFP control for failed/moved/deleted service functions become more crucial in virtualized environments (e.g. ETSI NFV), where service functions are instantiated as VMs on servers. There is higher chance of state changes for those Service functions as the result of being decommissioned or replaced when over-utilized.

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

This draft uses the following terminologies defined by SFC-arch.

RSP: Rendered Service Path [SRC-arch]

SF: Service Function [SFC-arch].

SFC: Service Function Chain [SFC-arch].

SFF: Service Function Forwarder [SFC-arch].

SFP: Service Function Path [SFC-arch].

Here are the terminologies specific for this draft:

VSFI: SFC Visible Service Function Instance.

SFIC: Service Function Instance Component. One service function (e.g. NAT44) could have two different service function instantiations, one that applies policy-set-A (NAT44-A) and other that applies policy-set-B (NAT44-B). There could be multiple "entities" of NAT44-B (e.g. one "entity" only has 10G capability), and many "entities" of NAT44-B. Each entity has its own unique address. The "entity" in this context is called "Service Function Instance Component" (SFIC).

Service Chain: The sequence of service functions, e.g. Chain#1 {s1, s4, s6}, Chain#2{s4, s7} at functional level. Also see the definition of "Service Function Chain" in [SFC-Problem].

Service Chain Instance Path: The actual Service Function Instance Components selected for a service chain.

VNF: Virtualized Network Function [NFV-Terminology].

4. Background

4.1. Multiple Entities of one Service Function

One service function (say, NAT44) could have two different service function instantiations, one that applies to policy-set-A (NAT44-A) and other that applies to policy-set-B (NAT44-B). There could be multiple "entities" of NAT44-A (e.g. one "entity" only has 10G capability), and many "entities" of NAT44-B. Each entity has its own unique address (or Locator in [SFC-Reduction]). The "Entity" in this context is called "Service Function Instance Component" (SFIC).

Identical SFICs could be attached to different Service Function Forwarder (SFF) nodes. It is also possible to have multiple identical SFICs attached to one Service Function Forwarder (SFF) node, especially in a Network Function Virtualization (NFV) environment where each SFIC is a virtual service function with limited capacity.

At the functional level, the order of service functions, e.g. Chain#1 {s1, s4, s6}, Chain#2{s4, s7}, is important, but very often which SFIC of the Service Function "s1" is selected for the Chain #1 is not.

Some SFICs are visible to Service Chain Path. Sometimes a collection of SFICs can appear as one single entity to the Service

Chain Path. When multiple SFICs are attached to one SFF, the collection of all those SFICs can appear as a single Service Function to the Service Chain Path. As described in Section 5.5 of [SFC-arch], the SFF can make local decision in choosing the SFIC among the collection of directly attached identical SFICs. The individual SFIC in this collection doesn't have to be visible to the SFP, the classifier, or orchestration.

It is also possible that multiple SFICs of one service function can be reached by different SFF nodes as depicted by Figure 5 of [SFC-arch].

For the ease of description, the term "Service Function Instance" is used in this draft to represent the identical SFICs that are visible to the SFP. The identical SFICs attached to different SFFs are obviously visible to SFP. But the identical SFICs attached to one SFF via different ports can be local to the SFF, i.e. not visible to the SFP.

4.2. Rendered Service Path (RSP)

[SFC-arch] defines RSP as the constrained specification of where packets assigned to a certain service chain must go.

RSP can be logically represented by an ordered sequence of SFF nodes [SFF-sequence] and an ordered sequence of SFs on each SFF of the list [SFF-SF-sequence].

RSP can also be SF-sequence without specifying which SFFs for the SFs.

The SFF-SF-sequence can be explicitly encoded in the SFC header for the SFP, or can be passed down, as "traffic steering policies", to the relevant SFF nodes.

4.2.1. SFF-sequence and SFF-SF-sequence representation

Logically, the SFF-sequence is represented by a list of SFF nodes. For a Chain sf2 -> sf3 -> sf4 over the network depicted by the Figure 5 of SFC-arch (shown below with some minor changes), one RSP could be for packets to traverse sf2 & sf3 attached to sff-a followed by the sf4 attached to SFF-c. The corresponding SFF-sequence for the RSP is [sff-a -> sff-c]. The corresponding SFF-SF-sequence is [(sff-a: sf2->sf3)-> (sff-c: sf4)].

The SFF-sequence and/or SFF-SF-sequence, e.g. {sff-a, sff-c}, can be explicitly encoded in the SFC header for the SFP.

Alternatively, the SFF-sequence and/or SFF-SF-sequence can be passed down, as "traffic steering policies", to the "sff-a" and "sff-c" nodes for the SFP. The traffic steering policies can be represented as "matching" & "action".

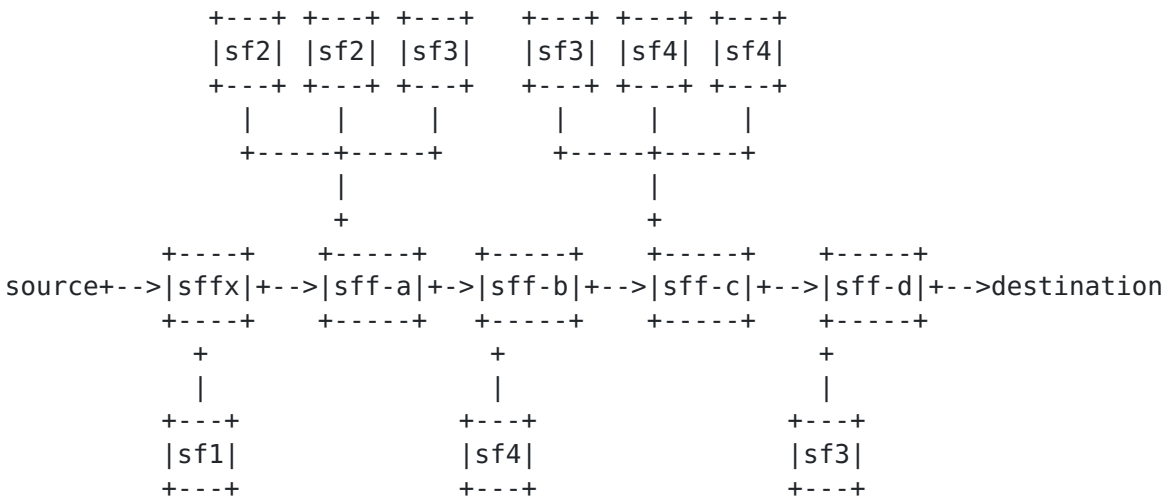


Figure 1:Service Function Attachment diagram

Suppose the SFC ID for this SFP is "yellow", the policy to "sff-a" can be:

Matching	Action
SFC ID="yellow"& ingress = sffx-port	next-hop: "sf2" &VID
SFC ID = "yellow" & ingress= sf2-port	next-hop: "sf3" &VID
SFC ID = "yellow" & ingress=sf3-port	next-hop: sff-b

Figure 2:Traffic Steering Policy to a SFF node

4.3. Multiple ways of Controlling RSP

How SFF-SF-sequence is selected for a given SFP to form the actual RSP is outside the scope of this draft. It is assumed that there is an external entity (e.g. service chain orchestration system) that is responsible for computing the SFF-sequence or SFF-SF-sequence for any given SFC.

This document focuses on the framework of replacing service functions for a given SFP/RSP.

To make the description easier, the following Service Chain architecture reference is used:

Some head end Service Chain Classifier can be configured with (or has the ability to specify) the exact SFF-SF-sequence for a given SFC. Some Classifier may only specify the SFF-sequence for a given SFC. Some Classifier may not specify SFF-sequence for a given SFC.

The SFF-SF-sequence or SFF-sequence can be

1. encoded in SFC header of every data packet;
2. Dynamic establishment of SFF-SF-sequence based on a SF-Sequence, which is almost like a list of IP addresses with each address representing one SF on the list; or
3. Dynamically programmed into relevant SFF nodes by a centralized network controller or a network management system, e.g. via I2RS interface.

The benefit of the Approach 1) above, i.e. encoding the exact path in every data packet, is no contention when there is change of RSP. The approach 1) above is basically "two dimensional" Source Routing, not only with explicit SFF nodes on the path, but also with exact SF sequence by each SFF node. Here are some issues associated with the Approach 1):

- For large flows, i.e. large number of packets in the flow, repeating the SFF-sequence/SFF-SF-sequence encoding in all packets may not be optimal, e.g. it can waste bandwidth which is not suitable for environment where bandwidth is limited.
- Whenever there is any state change to the SFs or SFFs on the path, the head end classification node has to be notified to encode a different path in data packets.

The approach 2) and 3) above are more appropriate for RSPs that don't change frequently. Not encoding the exact SFF-SF path in every data packet is beneficial to large flows.

When the in-band or out-of-band signaling methods are used to send the flow steering policies to the relevant SFF nodes, the packets associated with the SFP don't need to carry the SFF-SF-sequence or SFF-sequence. The forwarding nodes, e.g. SFFs, can establish the proper forwarding based on the steering policies. So they don't need to interpret the sequence carried by each packet.

The out-of-band method doesn't require the head end Service Chain Classifier to be configured with, nor has the capability to specify, the exact RSP. The out-of-band steering policies can be sent from an external entity, such as a centralized network controller or service chain orchestration system, e.g. via I2RS interface. Under this scenario, it doesn't require the head end Chain Classifier node to be aware of any change on the RSP.

There are times that it might not be feasible for the head end Service Chain Classifier to be notified of the changes of SFF-sequence or SFF-SF-Sequence for a given SFP because of the time taken for the notification and the limited capability of the Classifier nodes.

If each Service Function has a large number of SFICs, it scales better if the Classifier node doesn't need to be notified with status of SFICs on a SFP.

4.4. Impact of Virtualized Service Functions to SFP

When a SFP consists of virtualized service functions, e.g. in an ETSI NFV environment, the likelihood of changes to the corresponding RSP can be higher due to:

- Higher failure rate of virtualized service functions because most of them will not have build-in protection mechanism
- When a virtualized function is over-utilized, it is relatively easy to replace it by another one (SFIC) or instantiate more SFICs to take over the work load.

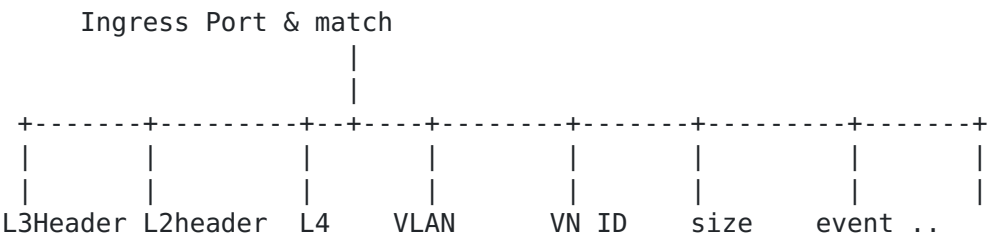
5. Steering Policies to SFF

It is assumed that there is an external service function chain manager or an orchestration system that computes the Service Function Path including the sequence of SFF nodes and the sequence of service functions for flows to traverse within each SFF node. It is beyond the scope of this draft on how the Service Function Chain orchestration system computes the path. This draft focuses on how & what the Service Function Orchestration pass to the Service Function Forwarder node on the specific policies, as shown in Figure below.

The SFF nodes are interconnected by tunnels, such as GRE, VxLAN, etc, and the SF are attached to a SFF node via Ethernet link or other link types. Therefore, the steering policies to a SFF node for service function chain depends on if the packet comes from previous SFF or comes from a specific SF. I.e. the SFC Service Layer Steering policies have to be ingress port specific. There are multiple different steering policies for one flow within one SFF and each set of steering policies is specific for an ingress port.

The semantics of traffic steering rules is "Match" and "Action", similar to the "route" described in [I-D.ietf-i2rs-rib-info-model]. The "match" & "action" for different ports can be different. The matching criteria for SFF can be more sophisticated. For example, the matching criteria could be any fields in the data packets:

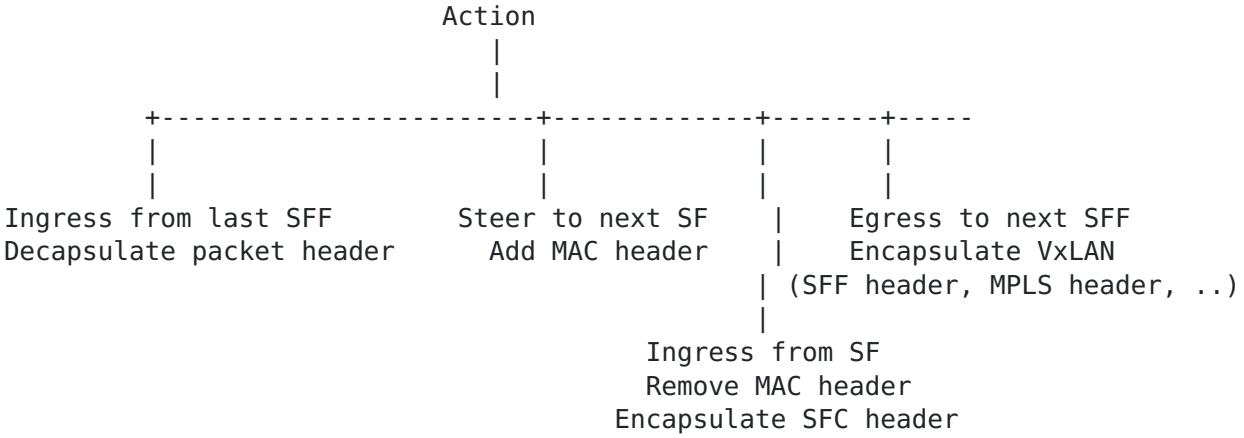
- Ingress port
- destination MAC,
- source MAC,
- VLAN_id,
- destination IP,
- source IP,
- TCP port,
- UDP port,
- QoS field,
- packet size, etc, or
- combination of any fields above.



A SFF node may not support some of the matching criteria listed above. It is important that Service Function Chain Orchestration System can retrieve the supported matching criteria by the SFF nodes.

The "Actions" for traffic steering could be to steer to the attached service function or instance via a specific port with specific VLAN-ID added, or next SFF nodes with specific VxLAN header.

When steering to the attached service function, the action has to include if additional VLAN-ID has to be added, or some header field of the packets have to be removed (for packets with certain header that is not supported by the attached service functions).



6. Local Restoration of Service Functions

When one SF Forwarder (SFF) node has multiple Service Function Instance Components (SFICs) of the same service function attached, the SFF can make a local decision on which SFIC is selected for a a given SFP, as described in Section 5.5 of [SFC-arch].

E.g. In the diagram below, The SF Forwarder (SFF) "A" has two instances of Service Function #7(SF7-1 & SF7-2), and 3 instances of Service Function #2 (SF2-2, SF2-4, SF2-5).

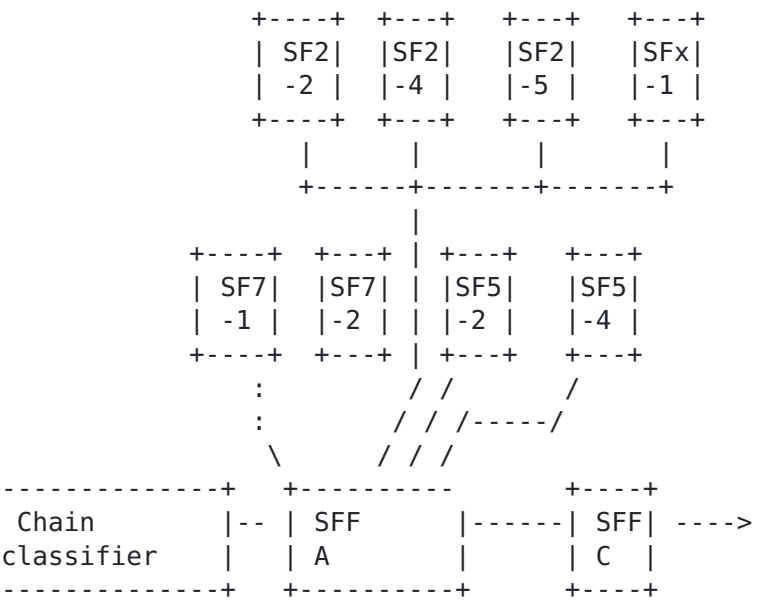


Figure 3:Local Restoration of Service Functions

For a service function chain that consists of "Service Function #7" followed by "Service Function #2", which is represented by SF7->SF2, the steering policy to SFF "A" could be simply SF7->SF2 without specifying which components of SF7 & SF2 are selected. In order for a SFF node to make local decision to choose one of the identical SFICs for a service function, the SFF node has to be aware of the SFICs for a given function on the SFP. The SFF node can be notified or configured with such information:

SF7 == {Port# for SF7-1, Port# for SF7-3}

SF2 == {Port# for SF2-2, Port# for SF2-4, Port# SF2-5}.

The multiple components within the {} represents the equal SFICs that the SFF can select locally.

The local protection and restoration is relatively simple and clean. ECMP can be used to balance all the available SFICs attached locally.

7. Global Restoration of Service functions

Sometimes changing the SFP's RSP involves using SFICs at different SFF nodes.

For a Chain sf2 -> sf3 -> sf4 in the Figure 5 of SFC-arch (with some minor changes):

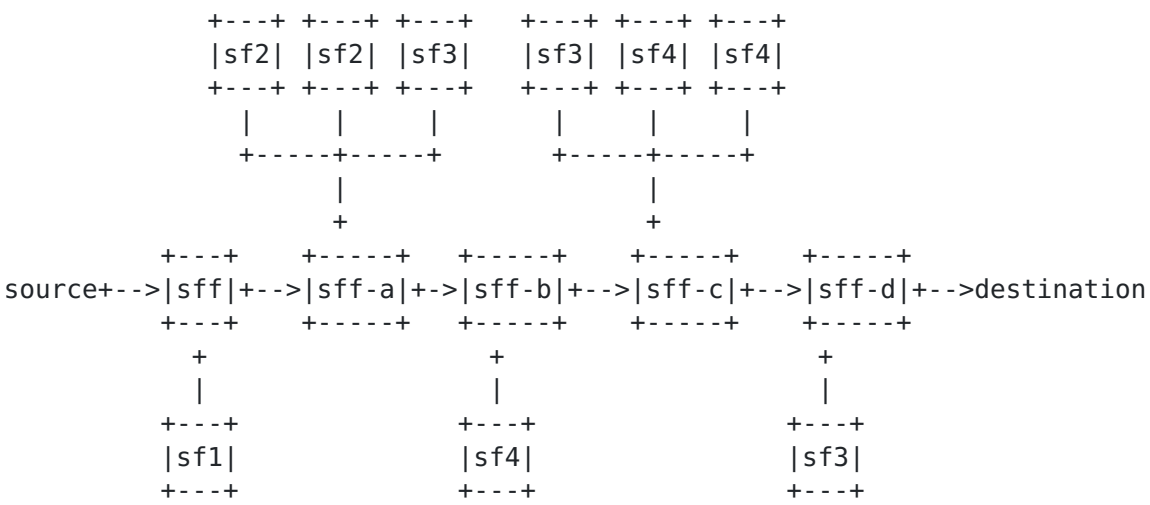


Figure 4: Global Restoration of Service Functions

Original Service Chain path: sf2 & sf3 at SFF-a; sf4 at SFF-c.

When the "sf3" attached to "sff-a" fails or over-utilized, the RSP needs to use the sf3 attached to "sff-c". The Path becomes:

- sf2 at "sff-a"; sf3 & sf4 at "sff-c".

This section examines possible ways to achieve the restoration when the change of SFP involves multiple SFF nodes.

7.1. Encoding the Exact SFF-SF-sequence in Data Packets

If the detailed SFF-SF-sequence is encoded in data packets, the SC Classifier needs to be notified of the changes of the RSP. The Classifier either gets notified of the exact SFF-SF-sequence from

external entity (e.g. controller or orchestration) or has the ability reconstruct the new RSP. The later approach needs protocol for the Classifier to be aware (or updated) of all the visible SFICs' states and their runtime topology.

Encoding the exact SFF-SF-sequence in every packet won't cause any contention issue among all the involved nodes when changes occur.

As mentioned in the previous section, encoding exact RSP path in every packet has the benefit and the issues of source routing. This approach may not be optimal when the RSP doesn't change very frequently, as in minutes or hours, or bandwidth is limited.

7.2. Dynamic establishment of an RSP

A similar method to MPLS RSVP-TE [RSVP-TE] signaling can be considered to dynamically establish the SFF-SF-sequence based on the SF-sequence.

Here is the overview of this approach. More details will be added later.

- The external controller computes the Service Chain Instance Path or Service Chain path at functional level and sent to the head end classifier node.
- The (segment) Head end Classifier node uses "Request for Path" signaling (like MPLS's RSVP) to establish the RSP to the nodes that on the path.
- All the nodes on the path establish the SF Forwarding Rule to the directly attached service functions (or the service function instances), and the appropriate tunnel from the egress port to the next SFF node for the given SFP.
- When the Path Confirmation is received (i.e. all the nodes along the path have completed the SF Forwarding Rule establishment and tunnel establishment), the head end can put user data along the pre-established Tunnel (e.g. VxLAN).

The drawback of this approach is that the head end node might receive packets belonging to the service function chain before all the involved nodes (SFF or SF) have made the needed changes.

It is very similar to the issues encountered by MPLS Fast Reroute [FRR]. MPLS FRR allows that packets to be dropped when a restoration path is being dynamically signaled because there was not a pre-established backup path.

7.3. Out-Of-Band Signaling of changes on SFP

If the out-of-band method is used, i.e. sending the updated flow steering policies to indicate the changes of the SFP path, there could be issues of synchronization and race conditions. For example, if the SFF "A" and SFF "C" get flow steering policies at slightly different times, some packets of the flow might miss some service functions on the chain.

In SDN or SDN-like environments, changes to a SFP can be dynamically programmed to relevant SFF nodes via out-of-band signal form a central controller or Network Management System (as in I2RS).

This approach does not require using end to end signaling protocol among Classier nodes and SFF nodes. But there may be problems introduced (such as loops or dropped packets) if SFF nodes are not updated in the proper order or not at the same time; the nodes should be updated in a similar time scale to the use of a signaling protocol. In addition, the network may have a single point of failure if the controller or NMS is not itself redundant.

7.4. Hybrid Method

For global restoration of service functions on a SFP, it is worthwhile to explore a hybrid mode, i.e. when there are changes involving using identical SFICs at different SFF nodes, the SC Classifier node is informed to encode the explicit SFFs for each SF in the SFC header of the data packets until all the involved SFF nodes complete the installation of the new steering policy for the path.

8. Regional Restoration of Service Function

It might not be always be feasible for the head end Service Chain Classifier to be aware of the exact SFICs selected for a given SFP due to too many SFICs for each SF, notifications not being promptly sent to the classifier node, or other reasons. Then Regional restoration should be considered.

This is not about multiple same-function SFICs attached to one SFF node. Those SFICs can be handled by the SFF via local load balance as described in SFC-Arch.

Regional restoration can take the similar approach as the Global restoration: choosing a regional ingress node that can take over the responsibility of installing the new steering policies to the involved SFF nodes or network nodes.

The Regional ingress node should be:

- on the data path of the flow of the given service chain;

- in front of the relevant the SFF nodes or network nodes that are impacted by the change of the Service Chain Path;

- capable of encoding the detailed Service Chain Path to the Service Chain Header of data packets of the identified flow; and

- capable of removing the detailed Service Chain Path encoding in data packets after all the impacted SFF nodes and network nodes completed the policy installation.

9. Conclusion and Recommendation

TBD

10. Manageability Considerations

TBD

11. Security Considerations

TBD

12. IANA Considerations

This document requires no IANA actions. RFC Editor: Please remove this section before publication.

13. References

13.1. Normative References

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

13.2. Informative References

[SFC-Problem] P. Quinn, et al, "Service Function Chaining Problem statement", draft-ietf-sfc-problem-statement-02, work in progress, April 2014

[NFV-Terminology] ETSI NFV ISG, "Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV", ETSI GS NFV 003 V1.1.1, Oct. 2013, http://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.01.01_60/gs_NFV003v010101p.pdf

[SFC-Reduction] R. Parker, "Service Function Chaining: Chain to Path Reduction", draft-parker-sfc-chain-to-path-00, work in progress, Nov. 2013

[RSVP-TE] D. Awduche, Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[FRR] P. Pan, Swallow, G., and Atlas, A., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005

14. Acknowledgments

Many thanks to Ron Bonica for the discussion in formulating the content for the draft.

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

Authors' Addresses

Linda Dunbar
Huawei Technologies
5340 Legacy Drive, Suite 175
Plano, TX 75024, USA
Phone: (469) 277 5840
Email: ldunbar@huawei.com

Andrew G. Malis
Huawei Technologies
USA
Email: agmalis@gmail.com