Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 23 August 2024

J. Dong
Z. Li
Huawei Technologies
C. Xie
C. Ma
China Telecom
G. Mishra
Verizon Inc.

20 February 2024

Carrying Network Resource Partition (NRP) Information in IPv6 Extension Header

draft-ietf-6man-enhanced-vpn-vtn-id-06

Abstract

Virtual Private Networks (VPNs) provide different customers with logically separated connectivity over a common network infrastructure. With the introduction and evolvement of 5G and also in some existing network scenarios, some customers may require network connectivity services with advanced features comparing to conventional VPN services. Such kind of network service is called enhanced VPNs. Enhanced VPNs can be used, for example, to deliver network slice services.

A Network Resource Partition (NRP) is a subset of the network resources and associated policies on each of a connected set of links in the underlay network. An NRP could be used as the underlay to support one or a group of enhanced VPN services. For packet forwarding in a specific NRP, some fields in the data packet are used to identify the NRP the packet belongs to, so that NRP-specific processing can be performed on each node along a path in the NRP.

This document specifies a new IPv6 Hop-by-Hop option to carry the NRP related information in data packets, which could be used to identify the NRP-specific processing to be performed on the packets by each network node along a network path in the NRP.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Table of Contents

1. Introduction

Virtual Private Networks (VPNs) [RFC4026] provide different customers with logically isolated connectivity over a common network infrastructure. With the introduction and evolvement of 5G and also in some existing network scenarios, some customers may require network connectivity services with advanced features comparing to conventional VPNs, such as resource isolation from other services or

guaranteed performance. Such kind of network service is called enhanced VPN [I-D.ietf-teas-enhanced-vpn]. Enhanced VPN service requires the coordination and integration between the overlay VPNs and the capability and resources of the underlay network. Enhanced VPN can be used, for example, to deliver IETF network slice services [I-D.ietf-teas-ietf-network-slices].

[I-D.ietf-teas-ietf-network-slices] also introduces the concept of the Network Resource Partition (NRP), which is a subset of the buffer/queuing/scheduling resources and associated policies on each of a connected set of links in the underlay network. An NRP can be associated with a logical network topology to select or specify the set of links and nodes involved.

[I-D.ietf-teas-enhanced-vpn] specifies the framework of NRP-based enhanced VPN and describes the candidate component technologies in different network planes and network layers. An NRP could be used as the underlay to meet the requirement of one or a group of enhanced VPN services.

In packet forwarding, traffic of different Enhanced VPN services needs to be processed separately based on the network resources and the logical topology associated with the corresponding NRP. [I-D.ietf-teas-nrp-scalability] describes the scalability considerations and the possible optimizations for providing a relatively large number of NRPs. One approach to improve the data plane scalability of NRP is to introduce a dedicated NRP ID in the data packet to identify the set of network resources allocated to an NRP, so that packets in an NRP can be processed and forwarded using the NRP-specific network resources, which could avoid possible resource competition with services in other NRPs. An NRP ID can have network resource semantics, which represents a subset of the resources (e.g. bandwidth, buffer and queuing resources) allocated on a given set of links and nodes which constitute a logical network topology. The logical topology of an NRP could be defined and identified using mechanisms such as Multi-Topology [RFC4915], [RFC5120] or Flex-Algo [RFC9350].

This document specifies a mechanism to carry NRP related information in a new IPv6 Hop-by-Hop option (Section 4.3 of [RFC8200]) called "NRP option". The NRP option is parsed by every intermediate node along the forwarding path, and the obtained NRP ID is used to invoke NRP-specific packet processing and forwarding using the set of NRP-specific resources. This provides a scalable solution to support a relatively large number of NRPs in an IPv6 network [I-D.ietf-teas-nrp-scalability].

Although in this document the application of the NRP option is to indicate the NRP-specific resource information, the NRP option is considered as a generic mechanism to convey network wide NRP ID and information with different semantics to meet the possible use cases in the future. Some considerations about generalization are described in Section 5.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. New IPv6 Extension Header Option for NRP

A new Hop-by-Hop option (Section 4.3 of [RFC8200]) type "NRP" is defined to carry the NRP related information. Its format is shown in Figure 1.

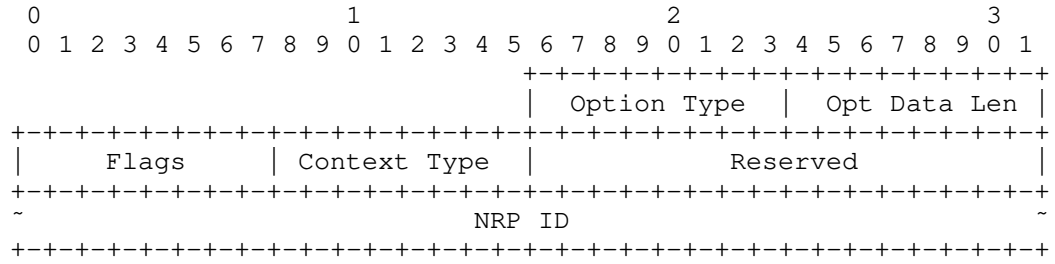


Figure 1. The format of NRP Option

Option Type: 8-bit identifier of the type of option. The type of NRP option is to be assigned by IANA. The bits of the type field are defined as below:

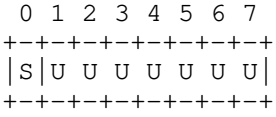
* BB 00 The highest-order 2 bits are set to 00 to indicate that a node which does not recognize this type will skip over it and continue processing the header.

* C 0 The third highest-order bit is set to 0 to indicate this option does not change en route.

* TTTTT To be assigned by IANA.

Opt Data Len: 8-bit unsigned integer indicates the length of the option Data field of this option, in octets.

Flags: 8-bit flags field. The most significant bit is defined in this document.



* S (Strict Match): The S flag is used to indicate whether the NRP ID MUST be strictly matched for the processing of the packet. It provides an approach for fine granular control of packet forwarding behavior when the NRP ID is not matched. If the NRP ID in the NRP option does not match with any of the NRP ID provisioned on the network node and the S flag is set to 1, the packet MUST be dropped. If the NRP ID does not match with any of the NRP ID provisioned on the network node and the S flag is set to 0, the packet MUST be forwarded using the default behavior as if the NRP option does not exist.

* U (Unused): These flags are reserved for future use. They MUST be set to 0 on transmission and MUST be ignored on receipt.

Context Type (CT): One-octet field used to indicate the semantics and length of the NRP ID carried in the option. The context value defined in this document is as follows:

* CT=0: The NRP ID is a 4-octet network-wide unique resource ID, which is used to identify the subset of network resources allocated to the NRP on the involved network nodes and links.

Reserved: 2-octet field reserved for future use. They MUST be set to 0 on transmission and MUST be ignored on receipt.

NRP ID: The identifier of a Network Resource Partition, the semantics and length of the ID is determined by the Context Type.

Note that, in the context of 5G network slicing, if a deployment found it useful, the four-octet NRP ID field may be derived from the four-octet Single Network Slice Selection Assistance Information (S-NSSAI) defined in 3GPP [TS23501].

3. Procedures

This section describes the procedures for NRP option processing when the Context Type in the NRP option is set to 0. The processing procedures for NRP option with other Context Types are out of the scope of this document and will be specified in separate documents which introduce those Context Types.

3.1. Adding NRP Option to Packets

When an ingress node of an IPv6 domain receives a packet, according to the traffic classification and mapping policy, the packet needs to be steered into one of the NRPs in the network, then the packet MUST be encapsulated in an outer IPv6 header with the source and destination addresses set according to the policy, and the NRP ID of the NRP which the packet is mapped to according to the policy MUST be carried in the NRP option of the Hop-by-Hop Options header, which is associated with the outer IPv6 header.

3.2. NRP-specific Packet Forwarding

On receipt of a packet with the NRP option, each network node which can process the Hop-by-Hop Options header and the NRP option in fast path [I-D.ietf-6man-hbh-processing] MUST use the NRP ID to determine the set of local network resources which are allocated to the NRP. The packet forwarding behavior is based on both the destination IP address and the NRP ID. More specifically, the destination IP address SHOULD be used to determine the next-hop and the outgoing interface, and NRP ID SHOULD be used to determine the set of network resources on the outgoing interface which are allocated to the NRP for processing and sending the packet. If the NRP ID does not match with any of the NRP ID provisioned on the outgoing interface, the S flag in the NRP option SHOULD be used to determine whether the packet should be dropped or forwarded using the default set of network resources of the outgoing interface. The Traffic Class field of the outer IPv6 header MAY be used to provide differentiated treatment for packets which belong to the same NRP. The egress node of the IPv6 domain MUST decapsulate the outer IPv6 header and the Hop-by-Hop Options header which includes the NRP option.

In the forwarding plane, there can be different approaches of partitioning the local network resources and allocating them to different NRPs. For example, on one physical interface, a subset of the forwarding plane resources (e.g. bandwidth and the associated buffer and queuing resources) can be allocated to a particular NRP and represented as a virtual sub-interface or a data channel with reserved bandwidth resource. In packet forwarding, the IPv6 destination address of the received packet is used to identify the next-hop and the outgoing layer-3 interface, and the NRP ID is used to further identify the virtual sub-interface or the data channel on the outgoing interface which is associated with the NRP.

Network nodes which do not support the processing of Hop-by-Hop Options header SHOULD ignore the Hop-by-Hop options header and forward the packet only based on the destination IP address. Network nodes which support Hop-by-Hop Options header, but do not support the

NRP option SHOULD ignore the NRP option and forward the packet only based on the destination IP address. The network node MAY process the rest of the Hop-by-Hop options in the Hop-by-Hop Options header.

4. Operational Considerations

As described in [RFC8200], network nodes may be configured to ignore the Hop-by-Hop Options header, drop packets containing a Hop-by-Hop Options header, or assign packets containing a Hop-by-Hop Options header to a slow processing path. In networks with such network nodes, it is important that packets of an NRP are not dropped due to the existence of the Hop-by-Hop Options header. Operators need to make sure that all the network nodes involved in an NRP can either process the Hop-by-Hop Options header in the fast path, or ignore the Hop-by-Hop Options header. Since an NRP is associated with a logical network topology, one practical approach is to ensure that all the network nodes involved in that logical topology support the processing of the Hop-by-Hop Options header and the NRP option in the fast path, and constrain the packet forwarding path to the logical topology of the NRP.

[I-D.ietf-6man-hbh-processing] specifies the modified procedures for the processing of IPv6 Hop-by-Hop Options header, with the purpose of making the Hop-by-Hop Options header useful. Network nodes complying with [I-D.ietf-6man-hbh-processing] will not drop packets with Hop-by-Hop Options header and the NRP option.

5. Considerations about Generalization

During the discussion of this document in the 6MAN WG, one of the suggestions received is to make the NRP option more generic in terms of semantics and encoding. This section gives some analysis about to what extent the semantics of NRP could be generalized, and how the generalization could be achieved with the proposed encoding.

Based on the NRP definition in [I-D.ietf-teas-ietf-network-slices], the concept of NRP could be extended as: an underlay network construct which is associated with a set of network-wide attributes and states maintained on each participating network node. The attributes associated with an NRP may include but not limited to: network resource attributes, network topology attributes, and network function attributes etc.

* The network resource can refer to various type of data plane resources, including link bandwidth, bufferage and queueing resources.

* The network topology can be multipoint-to-multipoint, point-to-point, point-to-multipoint or multipoint-to-point.

* The network functions may include both data forwarding actions and other types network functions which can be executed on data packets mapped to an NRP.

This shows the semantics of NRP can be quite generic. Although generalization is something good to have, it would be important to understand and identify the boundary of generalization. In this document, It is anticipated that for one network attribute to be included in NRP, it needs to be a network-wide attribute rather than a node-specific attribute. Thus whether a network-wide view can be provided or not could be considered as one prerequisite of making one attribute part of the NRP option.

The format of the NRP option contains the Flags field, the Context Type field and the Reserved field, which provide the capability for future extensions. That said, since the NRP option needs to be processed by network nodes in the fast path, the capability of network devices need to be considered when new semantics and encoding are introduced.

6. IANA Considerations

This document requests IANA to assign a new option type from "Destination Options and Hop-by-Hop Options" registry [IANA-HBH].

Hex Value	Binary Value act chg rest	Description	Reference
TBA	00 0 tba	NRP Option	[this document]

This document requests IANA to create a new registry for the "NRP Option Context Type" under the "Internet Protocol Version 6 (IPv6) Parameters" registry. The allocation policy of this registry is "Standards Action". The initial codepoints are assigned by this document as follows:

Value	Description	Reference
0	Resource ID	[this document]
1-254	Unassigned	
255	Reserved	[this document]

7. Security Considerations

The security considerations with IPv6 Hop-by-Hop Options header are described in [RFC8200], [RFC7045], [RFC9098] [RFC9099] and [I-D.ietf-6man-hbh-processing]. This document introduces a new IPv6 Hop-by-Hop option which is either processed in the fast path or ignored by network nodes, thus it does not introduce additional security issues.

8. Contributors

Zhibo Hu
Email: huzhibo@huawei.com

Lei Bao
Email: baolei7@huawei.com

9. Acknowledgements

The authors would like to thank Juhua Xu, James Guichard, Joel Halpern, Tom Petch, Aijun Wang, Zhenqiang Li, Tom Herbert, Adrian Farrel, Eric Vyncke, Erik Kline and Mohamed Boucadair for their review and valuable comments.

10. References

10.1. Normative References

[I-D.ietf-teas-enhanced-vpn]
Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A Framework for NRP-based Enhanced Virtual Private Network", Work in Progress, Internet-Draft, draft-ietf-teas-enhanced-vpn-17, 25 December 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-enhanced-vpn-17>.

[I-D.ietf-teas-ietf-network-slices]
Farrel, A., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slices-25, 14 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-25>.

[IANA-HBH] "IANA, "Destination Options and Hop-by-Hop Options"", 2016, <https://www.iana.org/assignments/ipv6-parameters/>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

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

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

10.2. Informative References

[I-D.ietf-6man-hbh-processing] Hinden, R. M. and G. Fairhurst, "IPv6 Hop-by-Hop Options Processing Procedures", Work in Progress, Internet-Draft, draft-ietf-6man-hbh-processing-13, 18 February 2024, <https://datatracker.ietf.org/api/v1/doc/document/draft-ietf-6man-hbh-processing/>.

[I-D.ietf-teas-nrp-scalability] Dong, J., Li, Z., Gong, L., Yang, G., Mishra, G. S., and F. Qin, "Scalability Considerations for Network Resource Partition", Work in Progress, Internet-Draft, draft-ietf-teas-nrp-scalability-03, 21 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-nrp-scalability-03>.

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, DOI 10.17487/RFC4026, March 2005, <https://www.rfc-editor.org/info/rfc4026>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC7045] Carpenter, B. and S. Jiang, "Transmission and Processing of IPv6 Extension Headers", RFC 7045, DOI 10.17487/RFC7045, December 2013, <https://www.rfc-editor.org/info/rfc7045>.

[RFC9098] Gont, F., Hilliard, N., Doering, G., Kumari, W., Huston, G., and W. Liu, "Operational Implications of IPv6 Packets with Extension Headers", RFC 9098, DOI 10.17487/RFC9098, September 2021, <https://www.rfc-editor.org/info/rfc9098>.

[RFC9099] Vyncke, É., Chittimaneni, K., Kaeo, M., and E. Rey, "Operational Security Considerations for IPv6 Networks", RFC 9099, DOI 10.17487/RFC9099, August 2021, <https://www.rfc-editor.org/info/rfc9099>.

[RFC9350] Psenak, P., Ed., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", RFC 9350, DOI 10.17487/RFC9350, February 2023, <https://www.rfc-editor.org/info/rfc9350>.

[TS23501] "3GPP TS23.501", 2016, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China
Email: jie.dong@huawei.com

Zhenbin Li
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China
Email: lizhenbin@huawei.com

Chongfeng Xie
China Telecom
China Telecom Beijing Information Science & Technology, Beiqijia
Beijing
102209
China
Email: xiechf@chinatelecom.cn

Chenhao Ma
China Telecom
China Telecom Beijing Information Science & Technology, Beiqijia
Beijing
102209
China
Email: machh@chinatelecom.cn

Gyan Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com

TEAS Working Group
Internet-Draft
Intended status: Informational
Expires: 5 September 2024

J. Dong
Z. Li
Huawei Technologies
L. Gong
China Mobile
G. Yang
China Telecom
G. Mishra
Verizon Inc.

4 March 2024

Scalability Considerations for Network Resource Partition
draft-ietf-teas-nrp-scalability-04

Abstract

A network slice offers connectivity services to a network slice customer with specific Service Level Objectives (SLOs) and Service Level Expectations (SLEs) over a common underlay network.

RFC XXXX describes a framework for network slices built using networks that use IETF technologies. As part of that framework, the Network Resource Partition (NRP) is introduced as a set of network resources that are allocated from the underlay network to carry a specific set of network slice service traffic and meet specific SLOs and SLEs.

As the demand for network slices increases, scalability becomes an important factor. Although the scalability of network slices can be improved by mapping a group of network slices to a single NRP, that design may not be suitable or possible for all deployments, thus there are concerns about the scalability of NRPs themselves.

This document discusses some considerations for NRP scalability in the control and data planes. It also investigates a set of optimization mechanisms.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Copyright Notice

Table of Contents

1. Introduction

RFC Editor Note: Please replace "RFC XXXX" in this document with the RFC number assigned to draft-ietf-teas-ietf-network-slices, and remove this note.

[I-D.ietf-teas-ietf-network-slices] defines network slicing in networks built using IETF technologies. These network slices may be referred to as RFC XXXX Network Slices, but in this document we simply use the term "network slice" to refer to this concept: this document only applies to the type of network slice described in [I-D.ietf-teas-ietf-network-slices].

The network slice aims to offer a connectivity service to a network slice customer with specific Service Level Objectives (SLOs) and Service Level Expectations (SLEs) over a common underlay network. [I-D.ietf-teas-ietf-network-slices] defines the terminologies and the characteristics of network slices. It also discusses the general framework, the components and interfaces for requesting and operating network slices. The concept of a Network Resource Partition (NRP) is introduced by [I-D.ietf-teas-ietf-network-slices] as part of the realization of network slices. An NRP is a collection of network resources in the underlay network, which can be used to ensure the requested SLOs and SLEs of network slice services are met.

[I-D.ietf-teas-enhanced-vpn] describes a layered architecture and the candidate technologies in different layers for delivering enhanced VPN services. Enhanced VPNs aim to meet the needs of customers or applications which require connectivity services with advanced characteristics, such as the assurance of SLOs and specific SLEs. Enhanced VPN services can be delivered by mapping one or a group of overlay VPNs to an NRP which is allocated with a set of network resources. The enhanced VPN architecture and technologies could be used for the realization of network slices.

As the demand for network slice services increases, scalability (the number of network slices a network can support within the capabilities and stabilities of the network protocols) becomes an important factor. Although the scalability of network slices can be improved by mapping a group of network slices to a single NRP, that design may not be suitable or possible for all deployments, thus there are concerns about the scalability of NRPs themselves.

This document discusses some considerations for NRP scalability in the control and data planes. It also investigates a set of optimization mechanisms.

2. Network Resource Partition Scalability Requirements

As described in [I-D.ietf-teas-ietf-network-slices], the connectivity constructs of network slices may be grouped together according to their characteristics (including SLOs and SLEs) and mapped to a given NRP. The grouping and mapping of network slices are policy-based and under the control of the network operator. For example, a network operator could consider a policy to host a large number of network slices using a relatively small number of NRPs to reduce the amount of state information to be maintained in the underlay network. On the other hand, a one-to-one mapping between network slices and NRPs gives more fine-grained control of the network slices, but comes at the cost of increased (per network slice) state in the underlay network.

With the introduction of various services that require enhanced connectivity, it is expected that the number of network slices will increase. The potential numbers of network slices and underlying NRPs are estimated by classifying the network slice deployment into three typical scenarios:

1. Network slices can be used by a network operator to deliver different types of service. For example, in a multi-service network, different network slices can be created to carry, e.g., mobile transport services, fixed broadband services, and enterprise services respectively. Each type of service could be managed by a separate team. Some other types of service, such as multicast services, may also be deployed in a separate virtual underlay network. A separate NRP may be created for each service type. It is also possible that a network infrastructure operator provides network slice services to other network operators as wholesale services, and an NRP may also be needed for each wholesale service operator. In this scenario, the number of NRPs in a network could be relatively small, such as in the order of 10 or so.

2. Network slice services can be requested by customers of industrial verticals, where the assurance of SLOs and the fulfilment of SLEs are contractually defined between the customer and the slice service provider, possibly including financial penalties if service provider fails to honor the contract. At the early stage of the vertical industrial deployment, a few customers in some industries will start using network slices to address the connectivity requirements and performance assurance

raised by their business, such as smart grid, manufacturing, public safety, on-line gaming, etc. The realization of such network slices may require the provision of different NRPs for different industries, and some customers may require dedicated NRPs for strict service performance guarantees. Considering the number of vertical industries and the number of customers in each industry, the number of NRPs needed may be in the order of 100.

3. With the advances in 5G and cloud networks, the type of network slices services defined in [I-D.ietf-teas-ietf-network-slices] could be widely used by customers of various vertical industries and enterprises who require guaranteed or predictable network service performance. The number of network slices in this case may increase to the order of thousands. Accordingly, the number of NRPs needed may be in the order of 1000.

In [TS23501], the 3GPP defines a 32-bit identifier for a 5G network slice with an 8-bit Slice/Service Type (SST) and a 24-bit Slice Differentiator (SD). This allows mobile networks (the Radio Access Networks (RANs) and mobile core networks) to potentially support a large number of 5G network slices. A 5G network slice is not the same as a network slice discussed in this document and defined in [I-D.ietf-teas-ietf-network-slices]. It is likely that multiple 5G network slices may be mapped to a single network slice defined by [I-D.ietf-teas-ietf-network-slices], but in some cases (for example, for specific SST or SD) the mapping may be closer to one-to-one. This may require increasing the number of network slices, the number of required NRPs may increase as well.

Thus the question of the scalability of network slice services arises. Mapping multiple network slices to a single NRP presents a significant scaling benefit, while a large number of NRPs may still be required, which raises scalability challenges too.

3. Scalability Design Principles

Scaling of network slicing can be achieved using a hierarchy of aggregation. Multiple network slices can be supported by a single NRP; multiple NRPs can be enabled on a filtered (logical) topology; and multiple filtered (logical) topologies utilise a single underlying network. The hierarchy, at any stage, may be made trivial (i.e., collapsed to a one-to-one mapping) according to the deployment objectives of the operator and the capabilities of the network technology.

To recap it in general terms:

* A network slice is an edge-to-edge service.

* An NRP is a set of network resources (e.g., buffers, bandwidth, queues) and the associated per-packet behaviors on a connected set of links in the underlay network.

* A filtered topology defines a collection of network links and nodes (call it a virtual network if it makes it easier for you to think about) on which path computation or traffic steering can be performed.

Scalability concerns exist at multiple points in the network slicing solution:

* The control protocols must be able to handle the distribution of information necessary to support the network slices, NRPs, and filtered topologies.

* The network nodes must be able to handle the computational load of determining paths based on the information of network slices, NRPs and filtered topologies..

* Path selection tools must be able to process network information and determine paths for network slice services on demand.

* The forwarding engines must be able to access the information in packets and make forwarding decisions at line speed.

Assuming that it is achievable, it is desirable for NRPs to have minimum impact (zero being preferred) on routing information that is propagated using IGP today, and to not require additional SPF computations beyond those that are necessary.

Assuming that an external mechanism can deal with path calculation and selection, it is desirable that in the calculated path information, the NRP identification should be decoupled from the information for path identification.

Given all of these considerations, we can set out the following design principles:

1. A filtered topology is a subset of the underlying physical topology. Thus, it defines which links (and nodes) are eligible to be used by the NRPs. It may be selected as a set of links with particular characteristics, or it may be a set of forwarding paradigms applied to the topology. Thus, a filtered topology may be realised through techniques such as multi-topology, coloring of links, virtual TE topology, or Flexible-Algorithm..

2. It is not envisaged that there would be many filtered topologies active, so running SPF per filtered topology is not a high burden.

3. Multiple NRPs can run on a single filtered topology meaning that the NRPs can be associated with the same filtered topology and use the SPF computation results from that topology.

4. Three separate things need to be identified by information carried within a packet:

   * Forwarding path (e.g. the next-hop)

   * NRP

   * Topology (i.e., filtered topology)

   How this information is encoded (using separate fields, same field, or overloading existing fields) forms part of the solution work.

5. NRP IDs should have domain-wide scope, and must be unique within a filtered topology.

6. Configuration mechanisms are used to set up packet/resource treatments on nodes.

7. Configuration mechanisms (such as southbound protocols from a controller) are used to set up resources and packet treatments of NRPs on the involved network nodes, and to install the bindings between domain-wide resource treatment identifiers (NRP IDs) and configured packet treatment.

8. The path computation or selection performed by or within a traffic engineering process, within or external to the head end node, (in particular the topology selection and path computation within that topology) may consider the characteristics of the filtered topology and the attributes of the NRP, but is agnostic to the resource treatment that the packets will receive within the network. Ensuring that the selected components of the path are capable of supporting the resource treatments identified by the NRP ID, is a separate matter.

9. The selected path is indicated in the packets using existing or new mechanisms. Whether that is SR-Policy, Flex-Algo, or something else is out of scope for this document, but it will obviously form part of the full set of network slicing solution specifications.

10. The components or mechanisms that are responsible for deciding what path to select for network slice service packet , for deciding how to mark the packets to follow the selected path, and for determining what resource treatment identifier (NRP ID) to apply to packets are also responsible for ensuring sufficient consistency so that the whole solution works.

Different operators can choose to deploy network slices at different scales, and while we may have opinions about what scales are sensible / workable / desirable, we do not attempt to constrain operators in their deployment choices.

The routing protocols (IGP or BGP) does not have to be involved in any of these points, but when they need to, it is important to isolate information of network slices and NRPs from existing routing information, so that there is no impact on scaling or stability. Furthermore, the complexity of SPF computation should not be impacted by the increasing number of network slices and NRPs.

Note that there is always a trade-off between optimal solutions and scalable solutions.

* We need to provide a scalable solution that can be deployed in all circumstances. We should acknowledge that:

  - We may need some extensions to the data/control/management plane to achieve this result. I.e., it may be that this cannot be done today with existing tools.

  - The scalable solution might not be optimal everywhere.

* We must understand that optimal solutions may be good for specific environments, but:

  - Might not work in some environments.

  - May have scalability issues in some other environments.

We should allow for both of these approaches, but we need to be clear of the costs and benefits in all cases in order that:

* We support significant optimisations and acknowledge the cost of necessary protocol extensions.

* We allow solutions which are suitable for specific environments, with their limitations documented so that they do not creep into wider deployment.

In particular, we should be open to the use of approaches that do not require control plane extensions and that can be applied to deployments with limited scope. Included in this are:

* Resource-aware SIDs

* L3VPN

It is anticipated that any specification of a network slicing protocol solution will include considerations of scalability and discussion of the applicability of the solution. This will not denegrate any specific solution, but will help clarify the type of deployment in which the solution is optimal while providing advice about its limitations in other deployments. Appendix A provides some simple examples of different possible realizations, and outlines their scaling properties.

4. Network Resource Partition Scalability Considerations

This section analyses the scalability of NRPs in the control plane and data plane to understand the possible gaps in meeting the scalability requirements.

4.1. Control Plane Scalability

The control plane for establishing and managing NRPs could be based on the combination of a centralized controller and a distributed control plane. The following subsections consider the scalability properties of both the distributed and the centralized control plane in such a design.

4.1.1. Distributed Control Plane

In some networks, multiple NRPs may need to be created for the delivery of network slice services. Each NRP is associated with a logical topology. The network resource attributes and the associated topology information of each NRP may need to be exchanged among the network nodes that participate in the NRP. The scalability of the distributed control plane used for the distribution of NRP information needs to be considered from the following aspects:

* The number of control protocol instances maintained on each node.

* The number of control protocol sessions maintained on each link.

* The number of control messages advertised by each node.

* The amount of attributes associated with each message (i.e., the size and the complexity of the messages).

* The number and frequency of computations (e.g., SPF computations) executed by each node.

As the number of NRPs increases, it is expected that, at least in some of the above aspects, the overhead in the control plane may increase in relation to the number of the NRPs. For example, the overhead of maintaining separate control protocol instances (e.g., IGP instances) for each NRP is considered higher than maintaining the information of multiple NRPs in the same control protocol instance with appropriate separation, and the overhead of maintaining separate protocol sessions for different NRPs is considered higher than using a shared protocol session for exchanging the information about multiple NRPs. To meet the scalability and performance requirements with the increasing number of NRPs, it is suggested to select the control plane mechanisms that have better scalability while still being able to provide the required functionality, isolation, and security for the NRPs.

4.1.2. Centralized Control Plane

The use of a centralized network controller may help to reduce the amount of computation overhead in the network, but may transfer some of the scalability concerns from network nodes to the network controller. Thus, the scalability of the controller also needs to be considered.

A centralized controller can have a global view of the network, and is usually used for Traffic Engineering (TE) path computation with various constraints, or for the global optimization of TE paths in the network. To provide TE path computation and optimization for multiple NRPs, the controller needs to know the up-to-date topology and resource information of all the NRPs. Additionally, for some events such as link or node failures, any updates to the NRPs may need to be distributed to the controller in real time, and may affect the planning and operation of some NRPs. When there is a significant change in the network which impacts multiple NRPs, or multiple NRPs require global optimization concurrently, there may be a heavy processing burden at the controller, and a large amount of signaling traffic to be exchanged between the controller and corresponding NRP components. These factors need to be taken into consideration from a scalability and performance standpoint.

4.2. Data Plane Scalability

Each NRP is allocated a subset of network resources. There may be a number of NRPs, each providing support for a set of network slices where each set of the slices requires a similar set of SLOs and SLEs. The sets of resources for each NRP may overlap, but may be independent to better enable the delivery of the SLOs and SLEs, and to avoid the risk of interference between the slices in different sets.

As the number of NRPs increases, the underlay network needs to provide a finer granularity of network resource partitioning, which means the amount of state maintained on the network nodes is likely to increase.

Network slice service traffic needs to be processed and forwarded by network nodes according to a forwarding policy that is associated with the topology and the resource attributes of the NRP it is mapped to, this means that some fields in the data packet need to be used to identify the NRP and its associated topology and resources either directly or implicitly. Different approaches for encapsulating the NRP information in data packets may have different scalability implications.

One practical approach is to reuse some of the existing fields in the data packet to indicate the NRP the packet belongs to. For example, the destination IP address or MPLS forwarding label could be reused to identify the NRP. This would avoid the complexity of introducing new fields into the data packet, but the additional semantics introduced to existing fields might require additional processing. Moreover, introducing NRP-specific semantics to existing fields in the packet could result in an increase in the number of identifiers (field values) that need to be assigned. Such an increase would be in proportion to the number of the NRPs. For example, if the destination IP address is used to identify an NRP, then a node which participates in M NRPs would need M IP addresses to be assigned to it. This might cause scalability problems in networks where a relatively large number of NRPs are in use.

An alternative approach is to introduce a new dedicated field in the data packet for identifying an NRP. And if this new field carries a network wide unique NRP identifier (NRP ID), it could be used together with the existing fields to determine the packet forwarding behavior. The potential issue with this approach lies in the difficulty of introducing a new field in some data plane technologies.

In addition, the introduction of NRP-specific packet forwarding impacts the number of the forwarding entries maintained by the network nodes.

5. Suggested Scalability Optimizations

To support more network slice services while keeping the amount of network state at a reasonable scale, one basic approach is to classify a set of network slice services (e.g., services which have similar service characteristics and performance requirements) into a group, and to map that group to an NRP, which is allocated with an aggregated set of network resources and the combination of the required logical topologies to meet the service requirements of the whole group of network slice services. Different groups of network slice services may be mapped to different NRPs, each of which is allocated with different set of network resources from the underlay network. According to the deployment policy of the operator, appropriate grouping of network slice services and mapping to NRPs could meet the network slice service requirements. However, in some network scenarios, such aggregation mechanism might not be applicable. The following sub-sections suggests further optimization in control plane and data plane respectively.

5.1. Control Plane Optimization

Control plane optimization may be considered in terms of distributed and centralized control planes.

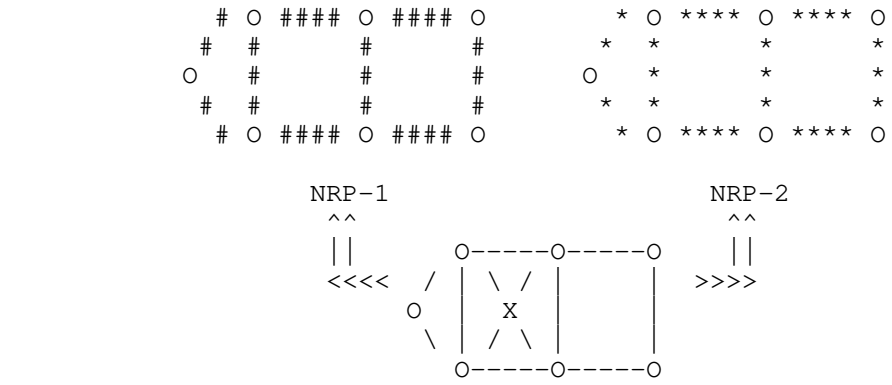
5.1.1. Distributed Control Plane Optimization

Several optimization mechanisms can be considered to reduce the distributed control plane overhead and improve its scalability.

The first control plane optimization consists of reducing the number of control plane sessions used for the establishment and maintenance of the NRPs. When multiple NRPs have the same connection relationship between two adjacent network nodes, an optimization could be achieved if a single control protocol session were used for all these NRPs. The information specific to the different NRPs could be exchanged over the same control protocol session, with necessary identifiers in the control messages to distinguish the information specific to different NRPs. This could reduce the overhead in a node of creating and maintaining a separate control protocol session for each NRP, and could also reduce the amount of control plane messages.

The second potential control plane optimization is to decouple the resource information of the NRP from the associated logical topology information, so that the resource attributes and the topology

attributes of the NRP can be advertised and processed separately. In a network, it is possible that multiple NRPs are associated with the same logical topology, or multiple NRPs may share the same set of network resources hosted by a specific set of network nodes and links. With topology sharing, it is more efficient to advertise only one copy of the topology information, and allow multiple NRPs deployed over the very same topology to exploit this topology information. More importantly, with this approach, the result of topology-based route computation can also be shared by multiple NRPs, so that the overhead of per NRP route computation is avoided. Similarly, for the resource sharing case, information about a set of network resources allocated on a particular network node or link could be advertised in the control plane only once, and then be referred to by multiple NRPs which share that set of resource.



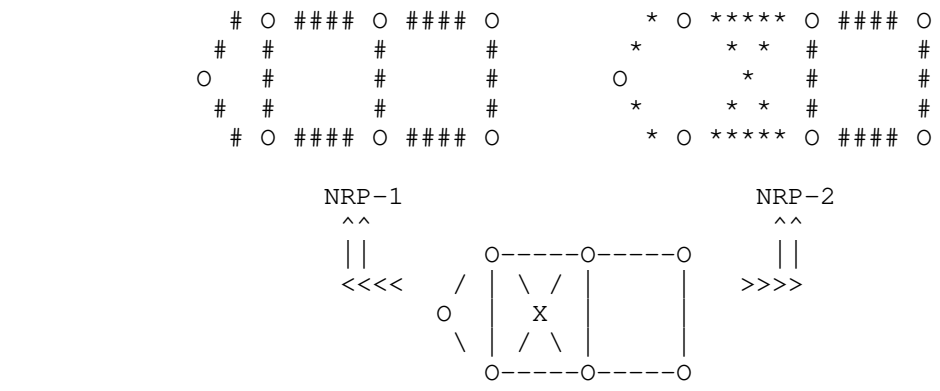
Underlay Network Topology

Legend

O Virtual node
Virtual links with a set of reserved resources
*** Virtual links with another set of reserved resources

Figure 1. Topology Sharing between NRPs

Figure 1 gives an example of two NRPs that share the same logical topology. NRP-1 and NRP-2 are associated with the same logical topology, while the resource attributes of each NRP are different. In this case, the information of the shared network topology can be advertised using either MT or Flex-Algo, then the two NRPs can be associated with the same MT or Flex-Algo, and the outcomes of topology-based route computation can be shared by the two NRPs for further generating the corresponding NRP-specific routing and forwarding entries.



Underlay Network Topology

Legend

O Virtual node
Virtual links with a set of reserved resource
*** Virtual links with another set of reserved resource

Figure 2. Resource Sharing between NRPs

Figure 2 gives another example of two NRPs which have different logical topologies, while they share the same set of network resources on a subset of the links. In this case, the information about the shared resources allocated on the those links needs to be advertised only once, and both NRP-1 and NRP-2 can refer to the common set of allocated link resources for constraint based path computation.

The control protocol extensions for support of scalable NRPs are out of the scope of this document. Proposals for solutions (such as that provided in [I-D.dong-lsr-sr-enhanced-vpn]) need to highlight their scalability properties and applicability so that implementers and deployers can make informed decisions.

5.1.2. Centralized Control Plane Optimization

For the optimization of the centralized control plane, it is suggested that the centralized controller is used to provide a computational facility to supplement the distributed control plane rather than as a replacement for the whole distributed control plane. In this way, the workload for NRP-specific path computation can be shared by both the centralized controller and the network nodes. In addition, the centralized controller may be realized through multiple network entities, each of which is responsible for one subset or

region of the network. This is the typical approach for scale-out of the centralized controller to avoid a single controller becoming congested or overloaded.

5.2. Data Plane Optimization

One potential optimization in the data plane consists of decoupling the identifiers used for topology-based forwarding from the identifier used for the NRP-inferred resource-specific processing. One possible mechanism is to introduce a dedicated network-wide NRP Identifier (NRP ID) in the packet header to uniquely identify the set of local network resources allocated to an NRP on each participating network node and link for the processing of packets. Then the existing identifiers in the packet header used for topology based forwarding (e.g., destination IP address, MPLS forwarding labels) are kept unchanged. The benefit is that the amount of the existing topology-specific identifiers will not be impacted by the increasing number of NRPs. Since this new NRP ID field will be used together with other existing fields of the packet to determine the packet forwarding behavior, this may require network nodes to maintain a hierarchical forwarding table in the data plane. Figure 3 shows the concept of using separate data plane identifiers for topology-specific and resource-specific packet forwarding and processing purposes.

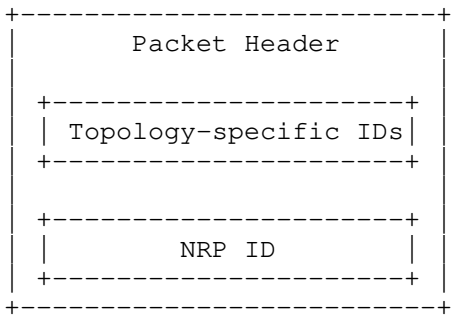


Figure 3. Decoupled Topology and Resource Identifiers in Data Packet

In an IPv6 [RFC8200] network, this could be achieved by carrying the NRP ID for the resource-specific forwarding in a new dedicated field in either the IPv6 base header or the extension headers. The destination IP address field would be retained for routing towards the destination prefix in the corresponding topology. Note that the NRP ID needs to be parsed by every node along the path that is capable of NRP-aware forwarding. [I-D.ietf-6man-enhanced-vpn-vtn-id] introduces a mechanism of carrying the NRP ID in the IPv6 Hop-by-Hop extension header.

In an MPLS [RFC3032] network, this may be achieved by inserting a dedicated NRP ID either in the MPLS label stack or as a specific field that follows the MPLS label stack. Thus, the existing MPLS forwarding labels are used for topology-specific packet forwarding purposes, and the NRP ID is used to determine the set of network resources for packet processing. This requires that both the forwarding label and the NRP ID are parsed by nodes along the forwarding path of the packet, and the forwarding behavior may depend on the position of the NRP ID in the packet. A possible approach for carrying the NRP ID is to use MPLS Network Actions (MNA) [I-D.ietf-mpls-mna-fwk], but specific solutions are out of the scope of this document.

6. Solution Evolution Perspectives

Based on the analysis provided by this document, the control and data plane for NRP may need to evolve to support the increasing number of network slice services and the increasing number of NRPs in the network. Appendix A provides some examples of network slicing solutions with limited applicability, and identifies their scalability concerns. However, a more generic and scalable solution could be more widely applicable and offer a future-proof mechanism. This section describes the evolution taking the SR-based NRP solutions as an example, while the analysis and optimization in this document are generic and not specific to SR.

First, by introducing resource-awareness with specific SR SIDs [I-D.ietf-spring-resource-aware-segments], and using Multi-Topology or Flex-Algo mechanisms to define the logical topology of the NRP, providing a small number of NRPs in the network is possible, and can meet the requirements for limited number of network slice services. This mechanism is called the "Basic SR-based NRP".

As the required number of network slice services increases, more NRPs may be needed, then the control plane scalability could be improved by decoupling the topology attributes from the resource attributes, so that multiple NRPs could share the same topology or resource attributes to reduce the overhead. The data plane can still rely on the resource-aware SIDs. This mechanism is called the "scalable SR-based NRP". Both the basic and the scalable SR-based NRP mechanisms are described in [I-D.ietf-spring-sr-for-enhanced-vpn].

Whenever the data plane scalability becomes a concern, a new dedicated NRP ID can be introduced in the data packet to decouple the resource-specific identifiers from the topology-specific identifiers in the data plane, so as to reduce the number of IP addresses or SR SIDs needed in supporting a large number of NRPs. This is called the NRP-ID-based mechanism.

7. Operational Considerations

The instantiation of NRPs require NRP-specific configurations of the participating network nodes and links. There can also be cases where the topology or the set of network resources allocated to an existing NRP needs to be modified. Of course, the amount of configurations for NRP instantiation and modification will increase with the number of NRPs.

For the management and operation of NRPs and the optimization of paths within the NRPs, the status of NRPs needs to be monitored and reported to the network controller. The increasing number of NRPs would require additional NRP status information to be monitored.

8. Security Considerations

This document discusses scalability considerations about the network control plane and data plane of NRPs in the realization of network slice services, and investigates some mechanisms for scalability optimization. As the number of NRPs supported in the data plane and control plane of the network can be limited, this may be exploited as an attack vector by requesting a large number of network slices, which then result in the creation of a large number of NRPs.

One protection against this is to improve the scalability of the system to support more NRPs. Another possible solution is to make the network slice controller aware of the scaling constraints of the system and dampen the arrival rate of new network slices and NRPs request, and raise alarms when the thresholds are crossed.

The security considerations in [I-D.ietf-teas-ietf-network-slices] and [I-D.ietf-teas-enhanced-vpn] also apply to this document.

9. IANA Considerations

This document makes no request of IANA.

10. Contributors

Fengwei Qin
qinfengwei@chinamobile.com

Jim Guichard
james.n.guichard@futurewei.com

Pavan Beeram
vbeeram@juniper.net

Tarek Saad
tsaad.net@gmail.com

Zhibo Hu
Email: huzhibo@huawei.com

Adrian Farrel
Email: adrian@olddog.co.uk

11. Acknowledgments

The authors would like to thank Adrian Farrel, Dhruv Dhody, Donald Eastlake, Kenichi Ogaki, Mohamed Boucadair, Christian Jacquenet and Kiran Makhijani for their review and valuable comments to this document.

Thanks, also, to the ad hoc design team of Les Ginsberg, Pavan Beeram, John Drake, Tarek Saad, Francois Clad, Tony Li, Adrian Farrel, Joel Halpern, and Peter Psenak who contributed substantially to establishing the design principles for scaling network slices.

12. References

12.1. Normative References

[I-D.ietf-teas-enhanced-vpn]
Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A Framework for NRP-based Enhanced Virtual Private Network", Work in Progress, Internet-Draft, draft-ietf-teas-enhanced-vpn-17, 25 December 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-enhanced-vpn-17>.

[I-D.ietf-teas-ietf-network-slices]
Farrel, A., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slices-25, 14 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-25>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

12.2. Informative References

[I-D.dong-lsr-sr-enhanced-vpn]
Dong, J., Hu, Z., Li, Z., Tang, X., Pang, R., and S. Bryant, "IGP Extensions for Scalable Segment Routing based Virtual Transport Network (VTN)", Work in Progress, Internet-Draft, draft-dong-lsr-sr-enhanced-vpn-10, 23 October 2023, <https://datatracker.ietf.org/doc/html/draft-dong-lsr-sr-enhanced-vpn-10>.

[I-D.ietf-6man-enhanced-vpn-vtn-id]
Dong, J., Li, Z., Xie, C., Ma, C., and G. S. Mishra, "Carrying Network Resource Partition (NRP) Information in IPv6 Extension Header", Work in Progress, Internet-Draft, draft-ietf-6man-enhanced-vpn-vtn-id-06, 20 February 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-6man-enhanced-vpn-vtn-id-06>.

[I-D.ietf-lsr-isis-sr-vtn-mt]
Xie, C., Ma, C., Dong, J., and Z. Li, "Applicability of IS-IS Multi-Topology (MT) for Segment Routing based Network Resource Partition (NRP)", Work in Progress, Internet-Draft, draft-ietf-lsr-isis-sr-vtn-mt-07, 23 January 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-isis-sr-vtn-mt-07>.

[I-D.ietf-mpls-mna-fwk]
Andersson, L., Bryant, S., Bocci, M., and T. Li, "MPLS Network Actions Framework", Work in Progress, Internet-

Draft, draft-ietf-mpls-mna-fwk-06, 24 January 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-mpls-mna-fwk-06>.

[I-D.ietf-spring-resource-aware-segments]
Dong, J., Miyasaka, T., Zhu, Y., Qin, F., and Z. Li, "Introducing Resource Awareness to SR Segments", Work in Progress, Internet-Draft, draft-ietf-spring-resource-aware-segments-08, 23 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-resource-aware-segments-08>.

[I-D.ietf-spring-sr-for-enhanced-vpn]
Dong, J., Miyasaka, T., Zhu, Y., Qin, F., and Z. Li, "Segment Routing based Network Resource Partition (NRP) for Enhanced VPN", Work in Progress, Internet-Draft, draft-ietf-spring-sr-for-enhanced-vpn-07, 4 March 2024, <https://datatracker.ietf.org/api/v1/doc/document/draft-ietf-spring-sr-for-enhanced-vpn/>.

[I-D.ietf-teas-ns-ip-mpls]
Saad, T., Beeram, V. P., Dong, J., Wen, B., Ceccarelli, D., Halpern, J. M., Peng, S., Chen, R., Liu, X., Contreras, L. M., Rokui, R., and L. Jalil, "Realizing Network Slices in IP/MPLS Networks", Work in Progress, Internet-Draft, draft-ietf-teas-ns-ip-mpls-03, 26 November 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ns-ip-mpls-03>.

[I-D.zhu-lsr-isis-sr-vtn-flexalgo]
Zhu, Y., Dong, J., and Z. Hu, "Using Flex-Algo for Segment Routing (SR) based Virtual Transport Network (VTN)", Work in Progress, Internet-Draft, draft-zhu-lsr-isis-sr-vtn-flexalgo-06, 10 July 2023, <https://datatracker.ietf.org/doc/html/draft-zhu-lsr-isis-sr-vtn-flexalgo-06>.

[RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, DOI 10.17487/RFC2702, September 1999, <https://www.rfc-editor.org/info/rfc2702>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC7926] Farrel, A., Ed., Drake, J., Bitar, N., Swallow, G., Ceccarelli, D., and X. Zhang, "Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks", BCP 206, RFC 7926, DOI 10.17487/RFC7926, July 2016, <https://www.rfc-editor.org/info/rfc7926>.

[RFC8453] Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for Abstraction and Control of TE Networks (ACTN)", RFC 8453, DOI 10.17487/RFC8453, August 2018, <https://www.rfc-editor.org/info/rfc8453>.

[RFC9350] Psenak, P., Ed., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", RFC 9350, DOI 10.17487/RFC9350, February 2023, <https://www.rfc-editor.org/info/rfc9350>.

[TS23501] "3GPP TS23.501", 2016, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>.

Appendix A. Example Network Slicing Realizations

This appendix contains some network slicing realisation examples. This is not intended to be a complete set of possible solutions, nor does it provide definitive ways to realize network slices and NRPs. The purpose of this appendix is to show how NRPs and network slices can be realised using existing tools and techniques, but explains some of the scaling issues that may arise and so how the applicability of these approaches may be limited. [I-D.ietf-teas-ns-ip-mpls] describes a scalable solution to realize network slicing in IP/MPLS networks.

A.1. VPNs with Default NRP

A possible deployment of network slices is to manage each network slice as one or multiple Layer-3 or Layer-2 VPNs and to support all of these VPNs on a single NRP that maps to the entire underlay network. This approach is perfectly functional and provides the required connectivity associated with the network slice services. In this case, the VPN identifiers may be considered as the identifiers of the network slices, which may present a scaling issue both in terms of the number of network slice identifiers available, and the routing protocols used for distribution of VPN routing information [RFC4364].

However, with only one NRP (the "default NRP"), the provision of services with sophisticated SLOs and SLEs may not be fulfilled, and additional network functions such as those discussed in Appendix A.3 and Appendix A.4 may be needed.

Thus, this approach may be a suitable solution for a modest number of network slices with relatively simple Service Level Agreements (SLAs), but more sophisticated approaches may be needed to address the scalability and advanced service levels with more than one NRPs.

A.2. Multiple Routing Instances for NRPs

An realization of NRP is to use a separate instance of the routing protocol for each independent network topology, and to map each topology to an NRP, which can be associated with a separate set of network links and nodes. The advantage of this approach is that each instance only has to handle advertisements for the links and nodes that are part of the topology, and for only one set of metrics.

However, each router that is in more than one topology must continue to run a SPF computations for each topology. It is possible that the routers do not need to maintain forwarding information for each topology if the destination addresses are assigned to specific topologies.

The biggest drawback is that routers that are part of more than one NRP must maintain separate protocol state for each protocol instance, and this may be a significant overhead. Further, run-time issues with one protocol instance may have a direct effect on the function of other protocol instances on the same router. Additionally, network operation and diagnostics may become complicated when protocol messages from multiple instances of a protocol are seen within the network.

Thus, this approach may be applicable only where NRPs use underlying topology resources that do not overlap significantly, and, in any case, only for a very small number of NRPs.

A.3. Resource-Aware Segment Routing based NRPs

One existing mechanism of building NRPs is to use resource-aware Segment Identifiers (either SR-MPLS or SRv6) [I-D.ietf-spring-resource-aware-segments] to identify the subset of network resources allocated to NRPs in the data plane based on the mechanisms described in [I-D.ietf-spring-sr-for-enhanced-vpn]. Network slices can be provisioned as L3 or L2 VPNs similar to the mechanisms described in Appendix A.1.This approach can provide NRPs with dedicated network resources, thus allows the SLOs and SLEs of the network slices to be met.

In the control plane, Multi-topology routing ([RFC4915] and [RFC5120]) can be used to define a small number of independent network topologies within the same routing protocol instance. Each topology can be associated with an NRP that supports a set of network slices. Alternatively, Flexible Algorithm [RFC9350] can be used to define the topological constraints and calculation algorithms for distributed path computation, which allows different forwarding paradigms to be constructed on the underlay network. So each NRP can also be assigned with a different Flex-Algo. The NRP resource attributes and the associated topology or topology constraints can be distributed using mechanisms based on Multi-topology [I-D.ietf-lsr-isis-sr-vtn-mt] or Flex-Algo [I-D.zhu-lsr-isis-sr-vtn-flexalgo].

It is suitable for networks where a relatively small number of NRPs are needed. As the number of NRPs increases, there may be several scalability challenges with this approach:

1. In OSPFv2, only 128 values are available to identify the different network topologies. In in IS-IS, 4096 values are available to identify different network topologies. As for Flex-Algo, only 128 unique Flex-Algo identifiers are available in the IGP extensions. This places a practical limit on the number of NRPs that can be supported in this approach.

2. A larger concern, however, is that SPF computations must be performed at each router for each topology. As the number of independent topologies increases, this computational load also increases and may become a burden for routers. This means that there may be a low limit to the number of NRPs that can be supported using this technique.

3. The number of resource-aware SR SIDs will increase in proportion to the number of NRPs, and the number of network segments (e.g. nodes and links) in the network, which will bring burden both to the distribution of the SR SIDs and related information in control protocols, and to the installation of forwarding table entries for those SIDs in the data plane.

A.4. MPLS-TE Virtual Networks

MPLS Traffic Engineering (MPLS-TE) ([RFC2702]) allows control of forwarding and the use of resources within an MPLS network. The use of resource reservation and the establishment of a set of traffic engineered MPLS label-switched paths (TE-LSPs) allows an MPLS network to be partitioned into multiple virtual networks ([RFC7926], [RFC8453]).

Each TE virtual network may be mapped to an NRP that supports a set of network slices. This can give a high level predictability to the NRP and allows the SLOs and SLEs of the network slices to be met.

However, each LSP must be planned, established, and maintained in the network. While this could be done using a central controller, it is usually achieved using the RSVP-TE signaling protocol [RFC3209]. Concerns have been expressed about the scalability of this protocol because it is a 'soft state' protocol, and because it requires a relatively large amount of state to be maintained on network nodes. Further, each virtual network (i.e., each NRP) requires a separate set of TE-LSPs meaning that the problem is not just O(n^2) for the mesh of LSPs between n edge nodes, but O(m*n^2) where there are m NRPs.

Thus, while this approach may facilitate high quality NRPs, it could present significant scaling concerns for the protocol engines on the routers in the network.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China
Email: jie.dong@huawei.com

Zhenbin Li
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China
Email: lizhenbin@huawei.com

Liyan Gong
China Mobile
No. 32 Xuanwumenxi Ave., Xicheng District
Beijing
China
Email: gongliyan@chinamobile.com

Guangming Yang
China Telecom
No.109 West Zhongshan Ave., Tianhe District
Guangzhou
China
Email: yangguangm@chinatelecom.cn

Gyan Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 17 September 2024

B. Wu
D. Dhody
Huawei Technologies
V.P. Beeram
Juniper Networks
T. Saad
Cisco Systems
S. Peng
ZTE Corporation

16 March 2024

YANG Data Models for Network Resource Partitions (NRPs)
draft-ietf-teas-nrp-yang-01

Abstract

RFC 9543 describes a framework for Network Slice in a network built from IETF technologies. In this framework, the network resource partition (NRP) is introduced as a collection of network resources allocated from the underlay network to carry a specific set of network slice service traffic and meet specific Service Level Objective (SLO) and Service Level Expectation (SLE) characteristics.

This document defines YANG data models for Network Resource Partitions (NRPs), applicable to devices and network controllers. The models can be used, in particular, for the realization of the RFC9543 Network Slice Services in IP/MPLS and Segment Routing (SR) networks.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 17 September 2024.

Table of Contents

1. Introduction

[RFC9543] describes a framework for Network Slice in a network built from IETF technologies. As specified in Section 7.4 [RFC9543], an NRP is a collection of resources identified in the underlay network to support the RFC9543 Network Slice Service to meet the slice Service Level Objectives (SLOs) and Service Level Expectations (SLEs) characteristics and network scalability.

This document defines two YANG models: NRPs network model in Section 4 and NRPs device model in Section 5. An Network Slice Controller (NSC) defined in Section 6.3 [RFC9543] can use the NRP network model to manage NRP instances for Network Slice Services. According to the YANG model classification of [RFC8309], the NRPs network model is a network configuration model.The NRPs device model can be used for device configuration, including device-specific configuration (e.g. interfaces).

The NRPs models conforms to the Network Management Datastore Architecture (NMDA) [RFC8342].

An NRP Policy [I-D.ietf-teas-ns-ip-mpls] is a policy construct that enables instantiation of mechanisms in support of service specific control and data plane behaviors on select topological elements associated with the NRP.Section 3.1 describes the detailed definition of NRP policy in NRP instantiation.

2. Terminology

The following terms are defined in [RFC6241] and are used in this specification:

* configuration data

* state data

The following terms are defined in [RFC7950] and are used in this specification:

* augment

* data model

* data node

The terminology for describing YANG data models is found in [RFC7950].

The tree diagram used in this document follows the notation defined in [RFC8340].

3. NRPs Data Model

The general operations of NRPs are as follows:

* NRPs instantiation: Depending on the slice service types and also network status, there can be two types of approaches. One method is to create an NRP instance before the network controller processes the Network Slice service request. Another one is that the network controller may start creating an NRP instance while configuring the Network Slice service request.

* NRPs modification: When the capacity of an existing NPR link is close to capacity, the bandwidth of the link could be increased. And when an NRP links or nodes resources are insufficient, new NRP links and nodes could be added.

* NRPs Deletion: If the NSC determines that no slice service is using an NRP, the NSC can delete the NRP instance.

* NRPs Monitoring: The NSC can use the NRPs model to track and monitor NRPs resource status and usage.

3.1. NRPs Instantiation

Section 3.5 in [I-D.ietf-teas-ns-ip-mpls] introduces the NRP policy. An NRP policy specifies the rules for determining the topology associated with the NRP and dictates how an NRP can be realized in IP/MPLS/SR networks. Section 4 of [I-D.ietf-teas-ns-ip-mpls] also defines three partition modes: (a) just the data plane or in (b) just the control plane or in (c) both the control and data planes. The NRP policy can dictate if the partitioning of the shared network resources can be achievedthrough one of the modes.

The NRP policy modes (a) and (c) require the forwarding engine on each NRP capable node to identify the traffic belonging to a specific NRP and to apply the corresponding Per-Hop Behavior (PHB) or forwarding mechanism that determines the forwarding treatment of the packets belonging to the NRP. When catering to Network Slices, this NRP identification is referred to as the NRP selector and may comprises of traffic streams from one or more connectivity constructs (belonging to one or more Network Slices) mapped to a specific NRP. The NRP policy modes (b) and (c) require the distributed/centralized resource reservation management.

'nrp-policy' is defined to enable NRP Stateful Traffic Engineering (NRP-TE) [I-D.ietf-teas-ns-ip-mpls] and/or NRP IGP forwarding in IP/MPLS networks [I-D.ietf-teas-nrp-scalability].

The high-level model structure of NRP policy defined by this document is as shown in Figure 1:

```
module: ietf-nrp
  augment /nw:networks:
    +--rw nrp-policies
       +--rw nrp-policy* [name]
          +--rw name                    string
          +--rw nrp-id?                 uint32
          +--rw mode?                   identityref
          +--rw resource-reservation
          |     ...
          +--rw selector
          |     ...
          +--rw phb-profile?            string
          +--rw topology
                ...
```

Figure 1: NRP Policy subtree high-level structure

The 'networks' container from the 'ietf-network' module [RFC8345] provides a placeholder for an inventory of nodes in the network. This container is augmented to carry a set of NRP policies.

The 'nrp-policies' container carries a list of NRP policies. Each 'nrp-policy' entry is identified by a name and holds the set of attributes needed to instantiate an NRP. Each entry also carries an 'nrp-id' leaf which uniquely identifies the NRP created by the enforcement of this policy.

The description of the 'nrp-policies' data nodes are as follows, and the other key elements of each nrp-policy entry are discussed in the following sub-sections.

* 'nrp-id': Is an identifier that is used to uniquely identify an NRP instance within an NSC network scope.

* 'mode': Refers to control plane resource partition, data plane resource partition, or a combination of both types.

3.1.1. Resource Reservation

The 'resource-reservation' container specifies the bandwidth resource allocated to an NRP instance, or can be overridden by the configuration of the link specific 'resource-reservation' nodes of 'nrp-topology'.

```
+--rw resource-reservation
  +--rw (max-bw-type)?
     +--:(bw-value)
     |  +--rw maximum-bandwidth?           uint64
     +--:(bw-percentage)
        +--rw maximum-bandwidth-percent?   rt-types:percentage
```

Figure 2: NRP Resource Reservation YANG subtree structure

3.1.2. NRP Selector

NRP selector defines the data plane encapsulation types and values that are used to identify NRP-specific network resources.

[I-D.ietf-teas-nrp-scalability] discusses several candidate NRP selector encapsulation schemes, including IP, MPLS, or SRv6, for example, the IPv6 Hop-by-Hop extension header defined in [I-D.ietf-6man-enhanced-vpn-vtn-id], or the SRv6 SID defined in [I-D.ietf-spring-sr-for-enhanced-vpn]. Since the MPLS encapsulation schemes are still under discussion, the model only provides a place holder for future updates. Additionally, the use of NRP-specific IP addresses to identify NRP resources, or the use of specific ACLs, are optional NRP selector mechanisms.

```
+--rw selector
|  +--rw ipv4
|  |  +--rw destination-prefix*   inet:ipv4-prefix
|  +--rw ipv6
|  |  +--rw (selector-type)?
|  |     +--:(dedicated)
|  |     |  +--rw ipv6-hbh-eh?          uint32
|  |     +--:(srv6-sid-derived)
|  |     |  +--rw srv6-sid*             inet:ipv6-prefix
|  |     +--:(ipv6-destination-derived)
|  |        +--rw destination-prefix*   inet:ipv6-prefix
|  +--rw mpls
|  +--rw acl-ref*   nrp-acl-ref
```

Figure 3: NRP Selector YANG subtree structure

3.1.3. Per-Hop Behavior (PHB)

PHB and NRP selector are combined mechanisms. PHB is used to specify the forwarding treatment of packets belonging to a specific NRP selector, such as bandwidth control, congestion control (e.g., Section 3.4 [RFC3644]). The exact definition of PHB is locally defined by the device or controller managing the NRPs. The 'phb-profile' leaf carries a name of a PHB profile available on the topological element where the policy is being enforced. Some examples of "phb-probile" may be standard PHBs, such as "Assured Forwarding (AF)", "Expedited Forwarding (EF)", or a customized local policies, such as "High", "Low", "Standard".

```
+--rw phb-profile?            string
```

Figure 4: PHB YANG subtree structure

3.1.4. NRP Topology

'nrp-topology' defines a dedicated NRP topology.

When an NRP support IGP forwarding, the topology of the NRP must be congruent with an IGP instance.The topology used for IGP route computation and forwarding can be derived using Multi-Topology Routing (MTR) or Flex-algo. Multi-Topology Routing (MTR) is defined in [RFC4915], [RFC5120], and [I-D.ietf-lsr-isis-sr-vtn-mt] or Flex-algo is defined in [RFC9350].

Figure 5 shows an example of NRP-1 enabling "igp-congruent", which indicates that this NRP instance uses the same IGP topology with the specified 'multi-topology-id' or 'algo-id'. As illustrated, NRP-1 has different link resource attributes from those of the IGP, but shares the same the nodes and termination point (TPs) of the IGP topology.

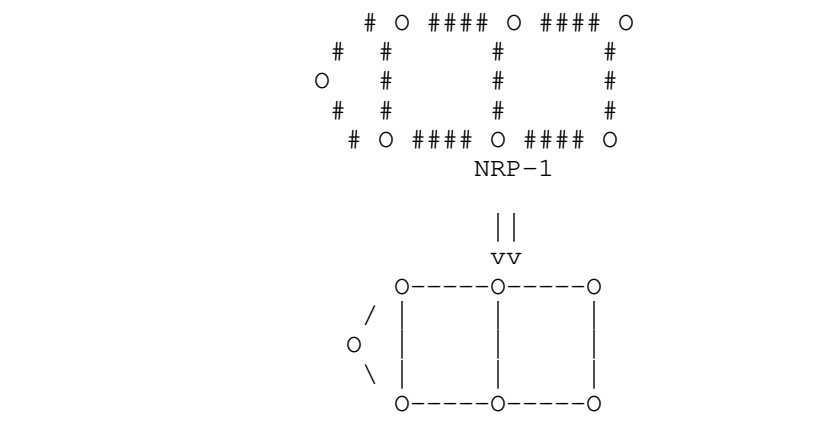


Figure 5: IGP Congruency Example

The 'selection' container consists of a list of select subset of links of an underlay topology or a pre-built topology.

The 'filter' container consists of a list of filters where each entry references a topology filter [I-D.bestbar-teas-yang-topology-filter]. The topological elements that satisfy the membership criteria can optionally override the default resource-reservation and nrp-selector specific leafs.

```
   +--rw topology
      +--rw igp-congruent!
      |  +--rw multi-topology-id?   uint32
      |  +--rw algo-id?             uint32
      |  +--rw sharing?             boolean
      +--rw (topology-type)?
         +--:(selection)
         |  +--rw select
         |     +--rw topology-group* [group-id]
         |        +--rw group-id                string
         |        +--rw base-topology-ref
         |        |     ...
         |        +--rw links* [link-ref]
         |        |     ...
         |        +--rw resource-reservation
         |        |     ...
         |        +--rw link-partition-type?
         |        |       identityref
         |        +--rw phb-profile?            string
         +--:(filter)
            +--rw filters
               +--rw filter* [filter-ref]
                  +--rw filter-ref
                  |       nrp-topo-filter-ref
                  +--rw resource-reservation
                  |     ...
                  +--rw selector
                  |     ...
                  +--rw phb-profile?           string
```

Figure 6: NRP Topology YANG subtree structure

3.2. NRPs monitoring

The NRPs model can be used to track and monitor operational status and resource usage of NRPs.

```
augment /nw:networks/nw:network/nw:network-types:
  +--rw nrp!
augment /nw:networks/nw:network/nw:node:
  +--ro nrp
     +--ro nrp-aware-dp-id
           ...
augment /nw:networks/nw:network/nt:link:
  +--ro nrp
     +--ro link-partition-type?   identityref
     +--ro bandwidth-value?       uint64
     +--ro nrp-aware-dp-id
     |     ...
     +--ro statistics
           ...
augment /nw:networks/nw:network/nw:node:
  +--ro nrps* [nrp-id]
     +--ro nrp-id    uint32
     +--ro nrp
           ...
augment /nw:networks/nw:network/nt:link:
  +--ro nrps* [nrp-id]
     +--ro nrp-id                 uint32
     +--ro link-partition-type?   identityref
     +--ro bandwidth-value?       uint64
     +--ro nrp-aware-dp-id
           ...
```

Figure 7: NRPs Monitoring YANG subtree structure

3.3. NRPs Device Model Description

The device-specific NRPs model is defined in module 'ietf-nrp-device' as shown in Section 5, which augments NRPs YANG data model in Section 4 and adds interface attributes, including resource reservation, NRP selector, and PHB profile, that are specific to an NRP device.

Figure below shows the tree diagram of the device NRPs YANG model defined in modules 'ietf-nrp-device.yang'.

```
module: ietf-nrp-device
  augment /nw:networks/nrp:nrp-policies/nrp:nrp-policy:
    +--rw interfaces
       +--rw interface* [interface]
          +--rw interface               if:interface-ref
          +--rw resource-reservation
          |  +--rw (max-bw-type)?
          |     +--:(bw-value)
          |     |     ...
          |     +--:(bw-percentage)
          |           ...
          +--rw selector
          |  +--rw ipv4
          |  |  +--rw destination-prefix*   inet:ipv4-prefix
          |  +--rw ipv6
          |  |  +--rw (selector-type)?
          |  |        ...
          |  +--rw mpls
          |  |  +--rw (selector-type)?
          |  |        ...
          |  +--rw acl-ref*   nrp-acl-ref
          +--rw phb-profile?            string
```

Figure 8: NRPs Device YANG subtree high-level structure

4. NRPs Yang Module

The 'ietf-nrp' module uses types defined in [RFC8345], [RFC8294],[RFC8776], [RFC6991], [RFC8519], [I-D.ietf-spring-srv6-yang], and [I-D.bestbar-teas-yang-topology-filter].

<CODE BEGINS> file "ietf-nrp@2024-01-03.yang"

```
module ietf-nrp {
  yang-version 1.1;
  namespace "urn:ietf:params:xml:ns:yang:ietf-nrp";
  prefix nrp;

  import ietf-network {
    prefix nw;
    reference
      "RFC 8345: A YANG Data Model for Network Topologies";
  }
  import ietf-network-topology {
    prefix nt;
    reference
      "RFC 8345: A YANG Data Model for Network Topologies";
  }
```

```
  import ietf-routing-types {
    prefix rt-types;
    reference
      "RFC 8294: Common YANG Data Types for the Routing Area";
  }
  import ietf-te-types {
    prefix te-types;
    reference
      "RFC 8776: Traffic Engineering Common YANG Types";
  }
  import ietf-te-packet-types {
    prefix te-packet-types;
    reference
      "RFC 8776: Traffic Engineering Common YANG Types";
  }
  import ietf-inet-types {
    prefix inet;
    reference
      "RFC 6991: Common YANG Data Types";
  }
  import ietf-access-control-list {
    prefix acl;
    reference
      "RFC 8519: YANG Data Model for Network Access Control Lists
       (ACLs)";
  }
  import ietf-srv6-types {
    prefix srv6-types;
    reference
      "draft-ietf-spring-srv6-yang: YANG Data Model for SRv6 Base
       and Static";
  }
  import ietf-topology-filter {
    prefix topo-filt;
    reference
      "draft-bestbar-teas-yang-topology-filter: YANG Data Model
       for Topology Filter";
  }

  organization
    "IETF Traffic Engineering Architecture and Signaling (TEAS)
     Working Group";
  contact
    "WG Web:   <https://datatracker.ietf.org/wg/teas/>
     WG List:  <mailto:teas@ietf.org>

     Editor:   Bo Wu
            <mailto:lana.wubo@huawei.com>
```

```
        Editor:   Dhruv Dhody
                 <mailto:dhruv.ietf@gmail.com>

        Editor:   Vishnu Pavan Beeram
                  <mailto:vbeeram@juniper.net>

        Editor:   Tarek Saad
                  <mailto:tsaad.net@gmail.com>

        Editor:   Shaofu Peng
                  <mailto:peng.shaofu@zte.com.cn>";
    description
      "This YANG module defines a data model for
       Network Resource Partitions (NRPs) managment.

       Copyright (c) 2024 IETF Trust and the persons identified as
       authors of the code.  All rights reserved.

       Redistribution and use in source and binary forms, with or
       without modification, is permitted pursuant to, and subject
       to the license terms contained in, the Revised BSD License
       set forth in Section 4.c of the IETF Trust's Legal Provisions
       Relating to IETF Documents
          (https://trustee.ietf.org/license-info).

       This version of this YANG module is part of RFC XXXX
       (https://www.rfc-editor.org/info/rfcXXXX); see the RFC itself
       for full legal notices.";

    revision 2024-01-03 {
      description
        "Initial revision.";
      reference
        "RFC XXXX: A YANG Data Model for Network Resource
         Partitions (NRPs)";
    }

    /*
     * I D E N T I T I E S
     */

    identity nrp-partition-mode {
      description
        "Base identity for NRP partition type.";
    }

    identity nrp-control-plane-partition {
      base nrp-partition-mode;
```

```
    description
      "Identity for NRP control plane partition.";
  }

  identity nrp-data-plane-partition {
    base nrp-partition-mode;
    description
      "Identity for NRP data plane partition.";
  }

  identity nrp-hybrid-plane-partition {
    base nrp-partition-mode;
    description
      "Identity for both control and data planes partitions of NRP.";
  }

  identity nrp-link-partition-type {
    description
      "Base identity for NRP interface partition type.";
  }

  identity virtual-sub-interface-partition {
    base nrp-link-partition-type;
    description
      "Identity for NRP virtual interface or sub-interface partition,
       e.g. FlexE.";
  }

  identity queue-partition {
    base nrp-link-partition-type;
    description
      "Identity for NRP queue partition type.";
  }

  /*
   * T Y P E D E F S
   */

  typedef nrp-acl-ref {
    type leafref {
      path "/acl:acls/acl:acl/acl:name";
    }
    description
      "This type is used to reference an ACL.";
  }

  typedef nrp-topo-filter-ref {
    type leafref {
```

```
        path "/nw:networks/topo-filt:topology-filters/"
           + "topo-filt:topology-filter/topo-filt:name";
      }
      description
        "This type is used to reference a Topology Filter.";
      reference
        "draft-bestbar-teas-yang-topology-filter: YANG Data Model
         for Topology Filter";
    }

    /*
     * Grouping - NRP Resource Reservation
     */

    grouping nrp-resource-reservation {
      description
        "Grouping for NRP resource reservation.";
      container resource-reservation {
        description
          "Container for NRP resource reservation.";
        choice max-bw-type {
          description
            "Choice of maximum bandwidth specification.";
          case bw-value {
            leaf maximum-bandwidth {
              type uint64;
              units "bits/second";
              description
                "The maximum bandwidth allocated to an NRP
                 - specified as absolute value.";
            }
          }
          case bw-percentage {
            leaf maximum-bandwidth-percent {
              type rt-types:percentage;
              description
                "The maximum bandwidth allocated to an NRP
                 - specified as percentage of link
                 capacity.";
            }
          }
        }
      }
    }

    /*
     * Grouping - NRP Selector Configuration
     */
```

```
    grouping nrp-selector-config {
      description
        "Grouping for NRP selector configuration.";
      container selector {
        description
          "Container for NRP selector.";
        container ipv4 {
          description
            "Container for IPv4 NRP selector.";
          leaf-list destination-prefix {
            type inet:ipv4-prefix;
            description
              "Any prefix from the specified set of IPv4
               destination prefixes can be the selector.";
          }
        }
        container ipv6 {
          description
            "Container for IPv6 NRP selector.";
          choice selector-type {
            description
              "Choices for IPv6 selector type.";
            case dedicated {
              leaf ipv6-hbh-eh {
                type uint32;
                description
                  "The selector value carried in Hop-by-Hop
                   Option of IPv6 extension header.";
                reference
                  "draft-ietf-6man-enhanced-vpn-vtn-id: Carrying Virtual
                   Transport Network (VTN) Information in IPv6 Extension
                   Header";
              }
            }
            case srv6-sid-derived {
              leaf-list srv6-sid {
                type srv6-types:srv6-sid;
                description
                  "Any SID from the specified set of SRv6 SID can
                   be the selector.";
                reference
                  "draft-ietf-spring-sr-for-enhanced-vpn: Segment
                   Routing based Virtual Transport Network (VTN) for
                   Enhanced VPN";
              }
            }
            case ipv6-destination-derived {
              leaf-list destination-prefix {
```

```
                type inet:ipv6-prefix;
                description
                  "Any prefix from the specified set of IPv6
                   destination prefixes can be the selector.";
              }
            }
          }
        }
        container mpls {
          description
            "Container for MPLS NRP selector. This is a placeholder
             for future updates based on the MPLS solutions.";
        }
        leaf-list acl-ref {
          type nrp-acl-ref;
          description
            "Selection is done based on the specified list of ACLs.";
          reference
            "RFC 8519: YANG Data Model for Network Access Control Lists
             (ACLs)";
        }
      }
    }

    /*
     * Grouping - NRP QoS PHB profile
     */

    grouping nrp-qos-phb-profile {
      description
        "Grouping for NRP QoS PHB profile.";
      leaf phb-profile {
        type string;
        description
          "PHB profile identifier, specifying the forwarding treatment
           of packets belonging to a specific NRP selector, such as
           bandwidth control, congestion control
           (e.g., Section 3.4 [RFC3644]).  The PHB may be standard PHB,
           such as Assured Forwarding (AF), Expedited Forwarding (EF),
           or a customized local policy, such as 'High', 'Low',
           'Standard'.";
      }
    }

    /*
     * Grouping - NRP IGP congruent
     */
```

```
    grouping nrp-igp-congruent {
      description
        "Grouping for NRP IGP congruent attributes.";
      container igp-congruent {
        presence "Indicates NRP IGP congruency.";
        description
          "The presence of the container node describes NRP IGP
           congruent, which indicates that the NRP instance uses the same
           IGP topology with the specified 'multi-topology-id'
           and 'algo-id'. That is, the nodes and termination point of the
           NRP topology and the IGP topology are the same, while the link
           attributes of the NRP are different from those of the IGP.";
        leaf multi-topology-id {
          type uint32;
          description
            "Indicates the MT-id of the NRP IGP instance.";
          reference
            "RFC 5120:  M-ISIS: Multi Topology (MT) Routing in
             Intermediate System to Intermediate Systems (IS-ISs)
             RFC 4915:  Multi-Topology (MT) Routing in OSPF";
        }
        leaf algo-id {
          type uint32;
          description
            "Indicates the algo-id of the NRP IGP instance.";
          reference
            "RFC 9350: IGP Flexible Algorithm";
        }
        leaf sharing {
          type boolean;
          default "true";
          description
            "'true' if the the NRP IGP instance can be shared with
             other NRPs;
             'false' if the the NRP IGP instance is dedicated
             to this NRP.";
        }
      }
    }

    /*
     * Grouping - NRP Topology Filter
     */

    grouping nrp-topology-filter {
      description
        "Grouping for NRP filter topology.";
      container filters {
```

```
        description
          "Container for filters.";
        list filter {
          key "filter-ref";
          description
            "List of filters.";
          leaf filter-ref {
            type nrp-topo-filter-ref;
            description
              "Reference to a specific topology filter from the
               list of global topology filters.";
          }
          uses nrp-resource-reservation;
          uses nrp-selector-config;
          uses nrp-qos-phb-profile;
        }
      }
    }

    /*
     * Grouping - NRP Select Topology
     */

    grouping nrp-select-topology {
      description
        "NRP topology specified by selection.";
      container select {
        description
          "The container of NRP select topology.";
        list topology-group {
          key "group-id";
          description
            "List of groups for NRP topology elements (node or links)
             that share common attributes.";
          leaf group-id {
            type string;
            description
              "The NRP topology group identifier.";
          }
          container base-topology-ref {
            description
              "Container for the base topology reference.";
            uses nw:network-ref;
          }
          list links {
            key "link-ref";
            description
              "A list of links with common attributes";
```

```
            leaf link-ref {
              type leafref {
                path
                  "/nw:networks/nw:network[nw:network-id=current()"
                + "/../../base-topology-ref/network-ref]"
                + "/nt:link/nt:link-id";
              }
              description
                "A reference to a link in the base topology.";
            }
          }
          uses nrp-resource-reservation;
          leaf link-partition-type {
            type identityref {
              base nrp-link-partition-type;
            }
            description
              "Indicates the resource reservation type of an NRP link.";
          }
          uses nrp-qos-phb-profile;
        }
      }
    }

    /*
     * Grouping - NRP Topology
     */

    grouping nrp-topology {
      description
        "Grouping for NRP topology.";
      container topology {
        description
          "Container for NRP topology.";
        uses nrp-igp-congruent;
        choice topology-type {
          description
            "Choice of NRP topology type.";
          case selection {
            uses nrp-select-topology;
          }
          case filter {
            uses nrp-topology-filter;
          }
        }
      }
    }
```

```
     /*
      * Grouping - NRP Policy
      */

     grouping nrp-pol {
       description
         "Grouping for NRP policies.";
       container nrp-policies {
         description
           "Container for nrp policies.";
         list nrp-policy {
           key "name";
           unique "nrp-id";
           description
             "List of NRP policies.";
           leaf name {
             type string;
             description
               "A string that uniquely identifies the NRP policy.";
           }
           leaf nrp-id {
             type uint32;
             description
               "A 32-bit ID that uniquely identifies the NRP
                created by the enforcement of this NRP policy.";
           }
           leaf mode {
             type identityref {
               base nrp-partition-mode;
             }
             default "nrp-hybrid-plane-partition";
             description
               "Indicates the resource partition mode of the NRP, such as
                control plane partition, data plane partition,
                or hybrid partition.";
           }
           uses nrp-resource-reservation;
           uses nrp-selector-config;
           uses nrp-qos-phb-profile;
           uses nrp-topology;
         }
       }
     }

     /*
      * Grouping - NRP Selector State
      */
```

```
  grouping nrp-selector-state {
    description
      "The grouping of NRP selector.";
    container selector {
      config false;
      description
        "The container of NRP selector.";
      leaf srv6 {
        type srv6-types:srv6-sid;
        description
          "Indicates the SRv6 SID value as the NRP selector.";
      }
    }
  }

  /*
   * Grouping - NRP node attributes
   */

  grouping nrp-node-attributes {
    description
      "NRP node scope attributes.";
    container nrp {
      config false;
      description
        "Containing NRP attributes.";
      uses nrp-selector-state;
    }
  }

  /*
   * Grouping - NRP Link Attributes
   */

  grouping nrp-link-attributes {
    description
      "NRP link scope attributes.";
    leaf link-partition-type {
      type identityref {
        base nrp-link-partition-type;
      }
      config false;
      description
        "Indicates the resource partition type of an NRP link.";
    }
    leaf bandwidth-value {
      type uint64;
      units "bits/second";
```

```
        config false;
        description
          "Bandwidth allocation for the NRP as absolute value.";
      }
      uses nrp-selector-state;
    }

    /*
     * Grouping - NRP Bandwidth Metrics
     */

    grouping nrp-bandwidth-metrics {
      description
        "Grouping for NRP bandwidth metrics.";
      leaf one-way-available-bandwidth {
        type uint64;
        units "bits/second";
        description
          "Available bandwidth that is defined to be NRP link
           bandwidth minus bandwidth utilization..";
      }
      leaf one-way-utilized-bandwidth {
        type uint64;
        units "bits/second";
        description
          "Bandwidth utilization that represents the actual
           utilization of the link (i.e. as measured in the router).";
      }
    }

    // nrp-link-statistics

    grouping nrp-statistics-per-link {
      description
        "Statistics attributes per NRP link.";
      container statistics {
        config false;
        description
          "Statistics for NRP link.";
        leaf admin-status {
          type te-types:te-admin-status;
          description
            "The administrative state of the link.";
        }
        leaf oper-status {
          type te-types:te-oper-status;
          description
            "The current operational state of the link.";
```

```
        }
        uses nrp-bandwidth-metrics;
        uses te-packet-types:one-way-performance-metrics-packet;
      }
    }

    // nrp-network-type

    grouping nrp-network-type {
      description
        "Identifies the network type to be NRP.";
      container nrp {
        presence "Indicates NRP network topology.";
        description
          "The presence of the container node indicates NRP network.";
      }
    }

    /*
     * Augment - Network Resource Partition Policies.
     */

    augment "/nw:networks" {
      description
        "Augment networks with NRP policies.";
      uses nrp-pol;
    }

    /*
     * Augment - NRP type.
     */

    augment "/nw:networks/nw:network/nw:network-types" {
      description
        "Indicates the network type of NRP";
      uses nrp-network-type;
    }

    /*
     * Augment - NRP node operational status.
     */

    augment "/nw:networks/nw:network/nw:node" {
      when '../nw:network-types/nrp:nrp' {
        description
          "Augment only for NRP network topology.";
      }
      description
```

```
        "Augment node configuration and state.";
      uses nrp-node-attributes;
    }

    /*
     * Augment - NRP link operational status.
     */

    augment "/nw:networks/nw:network/nt:link" {
      when '../nw:network-types/nrp:nrp' {
        description
          "Augment only for NRP network topology.";
      }
      description
        "Augment link configuration and state.";
      container nrp {
        config false;
        description
          "Containing NRP attributes.";
        uses nrp-link-attributes;
        uses nrp-statistics-per-link;
      }
    }

    /*
     * Augment - Native topology with NRPs node operational status.
     */

    augment "/nw:networks/nw:network/nw:node" {
      description
        "Augment node with NRPs aware attributes.";
      list nrps {
        key "nrp-id";
        config false;
        description
          "List of NRPs.";
        leaf nrp-id {
          type uint32;
          description
            "NRP identifier";
        }
        uses nrp-node-attributes;
      }
    }

    /*
     * Augment - Native topology with NRPs link operational status.
     */
```

```
    augment "/nw:networks/nw:network/nt:link" {
      description
        "Augment link with NRPs aware attributes.";
      list nrps {
        key "nrp-id";
        config false;
        description
          "List of NRPs.";
        leaf nrp-id {
          type uint32;
          description
            "NRP identifier";
        }
        uses nrp-link-attributes;
      }
    }
  }
  <CODE ENDS>
```

Figure 9: NRPs data model YANG module

5. NRPs Device YANG Module

The device NRPs YANG module ('ietf-nrp-device') models augments the NRPs YANG module ('ietf-nrp') and adds the attributes of NRP interfaces that are local to an NRP device.

The device NRPs YANG module imports the following module(s): ietf-interfaces defined in [RFC8343], ietf-network defined in [RFC8345], and grouping defined in this document.

```
<CODE BEGINS> file "ietf-nrp-device@2024-01-03.yang"
module ietf-nrp-device {
  yang-version 1.1;
  namespace "urn:ietf:params:xml:ns:yang:ietf-nrp-device";
  prefix nrp-dev;

  /* Import IETF Network module */

  import ietf-network {
    prefix nw;
    reference
      "RFC 8345: RFC 8345: A YANG Data Model for Network Topologies";
  }

  /* Import IETF interface module */

  import ietf-interfaces {
```

```
      prefix if;
      reference
        "RFC8343: A YANG Data Model for Interface Management";
    }

    /* Import NRPs module */

    import ietf-nrp {
      prefix nrp;
      reference
        "RFCXXXX: A YANG Data Model for Network Resource
         Partitions (NRPs)";
    }

    organization
      "IETF Traffic Engineering Architecture and Signaling (TEAS)
       Working Group";
    contact
      "WG Web:   <https://datatracker.ietf.org/wg/teas/>
       WG List:  <mailto:teas@ietf.org>

       Editor:   Bo Wu
                 <mailto:lana.wubo@huawei.com>

       Editor:   Dhruv Dhody
                 <mailto:dhruv.ietf@gmail.com>

       Editor:   Vishnu Pavan Beeram
                 <mailto:vbeeram@juniper.net>

       Editor:   Tarek Saad
                 <mailto:tsaad.net@gmail.com>

       Editor:   Shaofu Peng
                 <mailto:peng.shaofu@zte.com.cn>";
    description
      "This YANG module defines a data model for Network Resource
       Partitions (NRPs) device configurations and states. The model
       fully conforms to the Network Management Datastore
       Architecture (NMDA).

       Copyright (c) 2024 IETF Trust and the persons identified as
       authors of the code.  All rights reserved.

       Redistribution and use in source and binary forms, with or
       without modification, is permitted pursuant to, and subject to
       the license terms contained in, the Revised BSD License set
       forth in Section 4.c of the IETF Trust's Legal Provisions
```

```
        Relating to IETF Documents
        (https://trustee.ietf.org/license-info).

        This version of this YANG module is part of RFC XXXX
        (https://www.rfc-editor.org/info/rfcXXXX); see the RFC itself
        for full legal notices.";

    // RFC Ed.: replace XXXX with actual RFC number and remove this
    // note.
    // RFC Ed.: update the date below with the date of RFC publication
    // and remove this note.

    revision 2024-01-03 {
      description
        "Initial revision.";
      reference
        "RFCXXXX: A YANG Data Model for Network Resource Partitions (NRPs)
         Device";
    }

    /* NRP device configuraiton */

    augment "/nw:networks/nrp:nrp-policies/nrp:nrp-policy" {
      description
        "NRP policy list.";
      /* NRP Interface Configuration Data */
      container interfaces {
        description
          "Configuration data model for NRP interfaces.";
        list interface {
          key "interface";
          description
            "NRP interfaces.";
          leaf interface {
            type if:interface-ref;
            description
              "NRP interface name.";
          }
          uses nrp:nrp-resource-reservation;
          uses nrp:nrp-selector-config;
          uses nrp:nrp-qos-phb-profile;
        }
      }
    }
  }
  <CODE ENDS>
```

Figure 10: NRPs Device data model YANG module

6. Scaling Considerations

[I-D.ietf-teas-nrp-scalability] analyzes the scalability considerations of the control plane and data plane in the NRPs deployment. This section complements some scalability considerations with the model and the possible implications on deployment or implementation.

Note: The possible managment impact of a large number of NRPs instance management on devices and controllers on a large-scale network scenarios will be added later.

7. Security Considerations

The YANG model defined in this document is designed to be accessed via network management protocols such as NETCONF [RFC6241] or RESTCONF [RFC8040]. The lowest NETCONF layer is the secure transport layer, and the mandatory-to-implement secure transport is Secure Shell (SSH) [RFC6242]. The lowest RESTCONF layer is HTTPS, and the mandatory-to-implement secure transport is TLS [RFC8446].

The NETCONF access control model [RFC8341] provides the means to restrict access for particular NETCONF or RESTCONF users to a preconfigured subset of all available NETCONF or RESTCONF protocol operations and content.

There are a number of data nodes defined in this YANG model that are writable/creatable/deletable (i.e., config true, which is the default). These data nodes may be considered sensitive or vulnerable in some network environments. Write operations (e.g., edit-config) to these data nodes without proper protection can have a negative effect on network operations.

nrp-link: A malicious client could attempt to remove a link from a topology, add a new link. In each case, the structure of the topology would be sabotaged, and this scenario could, for example, result in an NRP topology that is less than optimal.

The entries in the nodes above include the whole network configurations corresponding with the NRP, and indirectly create or modify the PE or P device configurations. Unexpected changes to these entries could lead to service disruption and/or network misbehavior.

8. IANA Considerations

This document registers a URI in the IETF XML registry [RFC3688]. Following the format in [RFC3688], the following registration is requested to be made:

```
URI: urn:ietf:params:xml:ns:yang:ietf-nrp
Registrant Contact: The IESG.
XML: N/A, the requested URI is an XML namespace.

URI: urn:ietf:params:xml:ns:yang:ietf-nrp-device
Registrant Contact: The IESG.
XML: N/A, the requested URI is an XML namespace.
```

This document requests to register a YANG module in the YANG Module Names registry [RFC7950].

```
Name: ietf-nrp
Namespace: urn:ietf:params:xml:ns:yang:ietf-nrp
Maintained by IANA: N
Prefix: nrp
Reference: RFC XXXX

Name: ietf-nrp-device
Namespace: urn:ietf:params:xml:ns:yang:ietf-nrp-device
Maintained by IANA: N
Prefix: nrp-dev
Reference: RFC XXXX
```

9. Acknowledgments

The authors would like to thank Krzysztof Szarkowicz, Jie Dong, Qin Wu, Yao Zhao, Zhenbing Li, Adrian Farrel, Tom Petch, Xuesong Geng, Italo Busi,and many others for their helpful comments and suggestions.

10. Contributor

The following individuals, authors of [I-D.bestbar-teas-yang-nrp-policy] and [I-D.wd-teas-nrp-yang], contributed to this consolidated document:

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

Mohamed Boucadair
Orange
Email: mohamed.boucadair@orange.com

Daniele Ceccarelli

Bin Wen
Comcast
Email: Bin_Wen@cable.comcast.com

Ran Chen
ZTE Corporation
Email: chen.ran@zte.com.cn

Luis M. Contreras
Telefonica
Email: luismiguel.contrerasmurillo@telefonica.com

Ying Cheng
China Unicom
Email: chengying10@chinaunicom.cn

Liyan Gong
China Mobile
Email: gongliyan@chinamobile.com

11. References

11.1. Normative References

[I-D.bestbar-teas-yang-topology-filter]
Beeram, V. P., Saad, T., Gandhi, R., and X. Liu, "YANG Data Model for Topology Filter", Work in Progress, Internet-Draft, draft-bestbar-teas-yang-topology-filter-05, 20 February 2024, <https://datatracker.ietf.org/doc/html/draft-bestbar-teas-yang-topology-filter-05>.

[I-D.ietf-6man-enhanced-vpn-vtn-id]
Dong, J., Li, Z., Xie, C., Ma, C., and G. S. Mishra, "Carrying Network Resource Partition (NRP) Information in IPv6 Extension Header", Work in Progress, Internet-Draft, draft-ietf-6man-enhanced-vpn-vtn-id-06, 20 February 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-6man-enhanced-vpn-vtn-id-06>.

[I-D.ietf-spring-srv6-yang]
Raza, S. K., Agarwal, S., Liu, X., Hu, Z., Hussain, I., Shah, H. C., Voyer, D., Elmalky, H., Matsushima, S., Horiba, K., Rajamanickam, J., and A. Abdelsalam, "YANG Data Model for SRv6 Base and Static", Work in Progress, Internet-Draft, draft-ietf-spring-srv6-yang-03, 4 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-srv6-yang-03>.

[RFC3688] Mealling, M., "The IETF XML Registry", BCP 81, RFC 3688, DOI 10.17487/RFC3688, January 2004, <https://www.rfc-editor.org/info/rfc3688>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure Shell (SSH)", RFC 6242, DOI 10.17487/RFC6242, June 2011, <https://www.rfc-editor.org/info/rfc6242>.

[RFC6991] Schoenwaelder, J., Ed., "Common YANG Data Types", RFC 6991, DOI 10.17487/RFC6991, July 2013, <https://www.rfc-editor.org/info/rfc6991>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/info/rfc7950>.

[RFC7951] Lhotka, L., "JSON Encoding of Data Modeled with YANG", RFC 7951, DOI 10.17487/RFC7951, August 2016, <https://www.rfc-editor.org/info/rfc7951>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[RFC8294] Liu, X., Qu, Y., Lindem, A., Hopps, C., and L. Berger, "Common YANG Data Types for the Routing Area", RFC 8294, DOI 10.17487/RFC8294, December 2017, <https://www.rfc-editor.org/info/rfc8294>.

[RFC8340] Bjorklund, M. and L. Berger, Ed., "YANG Tree Diagrams", BCP 215, RFC 8340, DOI 10.17487/RFC8340, March 2018, <https://www.rfc-editor.org/info/rfc8340>.

[RFC8341] Bierman, A. and M. Bjorklund, "Network Configuration Access Control Model", STD 91, RFC 8341, DOI 10.17487/RFC8341, March 2018, <https://www.rfc-editor.org/info/rfc8341>.

[RFC8342] Bjorklund, M., Schoenwaelder, J., Shafer, P., Watsen, K., and R. Wilton, "Network Management Datastore Architecture (NMDA)", RFC 8342, DOI 10.17487/RFC8342, March 2018, <https://www.rfc-editor.org/info/rfc8342>.

[RFC8343] Bjorklund, M., "A YANG Data Model for Interface Management", RFC 8343, DOI 10.17487/RFC8343, March 2018, <https://www.rfc-editor.org/info/rfc8343>.

[RFC8345] Clemm, A., Medved, J., Varga, R., Bahadur, N., Ananthakrishnan, H., and X. Liu, "A YANG Data Model for Network Topologies", RFC 8345, DOI 10.17487/RFC8345, March 2018, <https://www.rfc-editor.org/info/rfc8345>.

[RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/info/rfc8446>.

[RFC8519] Jethanandani, M., Agarwal, S., Huang, L., and D. Blair, "YANG Data Model for Network Access Control Lists (ACLs)", RFC 8519, DOI 10.17487/RFC8519, March 2019, <https://www.rfc-editor.org/info/rfc8519>.

[RFC8776] Saad, T., Gandhi, R., Liu, X., Beeram, V., and I. Bryskin, "Common YANG Data Types for Traffic Engineering", RFC 8776, DOI 10.17487/RFC8776, June 2020, <https://www.rfc-editor.org/info/rfc8776>.

[RFC9543] Farrel, A., Ed., Drake, J., Ed., Rokui, R., Homma, S., Makhijani, K., Contreras, L., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", RFC 9543, DOI 10.17487/RFC9543, March 2024, <https://www.rfc-editor.org/info/rfc9543>.

11.2. Informative References

[I-D.bestbar-teas-yang-nrp-policy]
Beeram, V. P., Saad, T., Wen, B., Ceccarelli, D., Peng, S., Chen, R., Contreras, L. M., and X. Liu, "YANG Data Model for Network Resource Partition Policy", Work in Progress, Internet-Draft, draft-bestbar-teas-yang-nrp-policy-03, 24 October 2022, <https://datatracker.ietf.org/doc/html/draft-bestbar-teas-yang-nrp-policy-03>.

[I-D.ietf-lsr-isis-sr-vtn-mt]
Xie, C., Ma, C., Dong, J., and Z. Li, "Applicability of IS-IS Multi-Topology (MT) for Segment Routing based Network Resource Partition (NRP)", Work in Progress, Internet-Draft, draft-ietf-lsr-isis-sr-vtn-mt-07, 23 January 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-isis-sr-vtn-mt-07>.

[I-D.ietf-spring-sr-for-enhanced-vpn]
Dong, J., Miyasaka, T., Zhu, Y., Qin, F., and Z. Li, "Segment Routing based Network Resource Partition (NRP) for Enhanced VPN", Work in Progress, Internet-Draft, draft-ietf-spring-sr-for-enhanced-vpn-07, 3 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-for-enhanced-vpn-07>.

[I-D.ietf-teas-nrp-scalability]
Dong, J., Li, Z., Gong, L., Yang, G., and G. S. Mishra, "Scalability Considerations for Network Resource Partition", Work in Progress, Internet-Draft, draft-ietf-teas-nrp-scalability-04, 4 March 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-nrp-scalability-04>.

[I-D.ietf-teas-ns-ip-mpls]
Saad, T., Beeram, V. P., Dong, J., Wen, B., Ceccarelli, D., Halpern, J. M., Peng, S., Chen, R., Liu, X., Contreras, L. M., Rokui, R., and L. Jalil, "Realizing Network Slices in IP/MPLS Networks", Work in Progress, Internet-Draft, draft-ietf-teas-ns-ip-mpls-03, 26 November 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ns-ip-mpls-03>.

[I-D.wd-teas-nrp-yang]
Wu, B., Dhody, D., Boucadair, M., Cheng, Y., and L. Gong, "A YANG Data Model for Network Resource Partitions (NRPs)", Work in Progress, Internet-Draft, draft-wd-teas-nrp-yang-02, 25 September 2022, <https://datatracker.ietf.org/doc/html/draft-wd-teas-nrp-yang-02>.

[RFC3644] Snir, Y., Ramberg, Y., Strassner, J., Cohen, R., and B. Moore, "Policy Quality of Service (QoS) Information Model", RFC 3644, DOI 10.17487/RFC3644, November 2003, <https://www.rfc-editor.org/info/rfc3644>.

[RFC8309] Wu, Q., Liu, W., and A. Farrel, "Service Models Explained", RFC 8309, DOI 10.17487/RFC8309, January 2018, <https://www.rfc-editor.org/info/rfc8309>.

[RFC9350] Psenak, P., Ed., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", RFC 9350, DOI 10.17487/RFC9350, February 2023, <https://www.rfc-editor.org/info/rfc9350>.

Appendix A. Open issues

This section lists the non-blocking issues raised during the Working Group adoption process. The issues listed below need to be fully resolved before publication

1. Raised by Tom Petch: Abstract lacks the reference to the NS framework that defines the NRP.

2. Raised by Adrain Farrel: 1) Avoid limiting IP/MPLS technology to realize NRPs, SR should be in scope; 2) Avoid the "IETF Network Slice" language, should use terms as "RFC 9543 Network Slice" and "RFC 9543 Network Slice Service" 3) It’s good to investigate any scaling issues with the model and any implications on deployments or implementations, just as draft-ietf-teas-nrp-scalability.

3. Raised by Med Boucadair: 1) Normative dependency on individual drafts, such as I-D.bestbar-teas-yang-topology-filter, suggesting to add it back when stable 2) The device model in the spec is not a device model as it augments a network model. 3) Some of the review provided in https://github.com/boucadair/IETF-Drafts-Reviews/blob/master/2024/draft-ahuang-netconf-udp-client-server-01-rev%20Med.pdf

4. Raised by Lius Contreras: 1) Clarify the NRP model usage in NSC, network controllers, and devices; 2) Rename Section 3.1.1 title to bandwidth reservation; 3) Add the references of "NRP capable node"; 4) In Section 3.1.3, better to clarify single PHB or multiple PHB per NRP and Whether the PHB management scope is in the NSC or network controller; 5) Section 3.1 adds description of NRP policy modes (b) and (c).

5. Raised by Xuesong: 1) Clarify the considerations for defining the NRP policy; 2) Distinguish NRP model operation and NRP mode (CP,DP, and hybrid); 3) Clarify the relationship and design consideration of NRPs network and device models.

6. Raised by Italo: 1) Clarify the models are technology-agnostic NRPs model or IP technology-specific NRPs model; 2) Updates the abstract/introduction to clarify that this model applies on devices and on controllers.

Appendix B. An Example

This section contains an example of an instance data tree in JSON encoding [RFC7951]. The example below instantiates an NRP for the topology that is depicted in the following diagram. There are three nodes, D1, D2, and D3. D1 has three termination points, 1-0-1, 1-2-1, and 1-3-1. D2 has three termination points as well, 2-1-1, 2-0-1, and 2-3-1. D3 has two termination points, 3-1-1 and 3-2-1. In addition there are six links, two between each pair of nodes with one going in each direction.

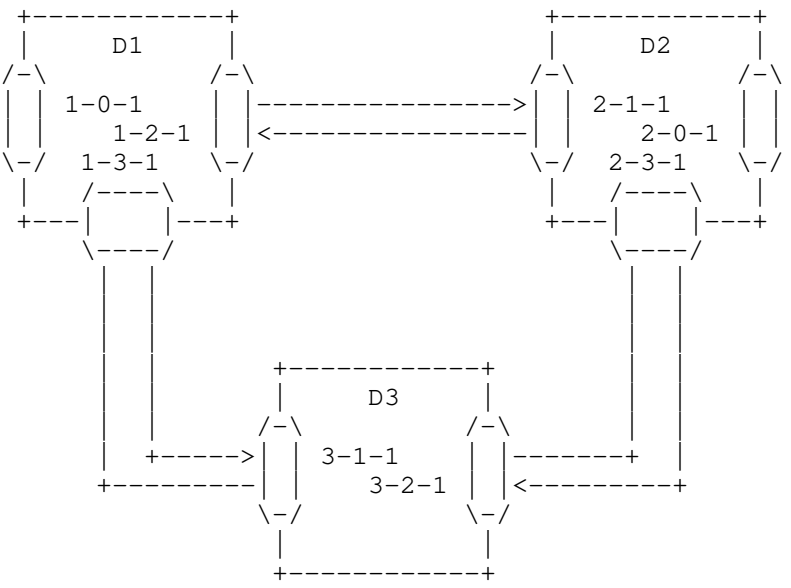


Figure 11: An NRP Instance Example

An corresponding IGP congruent NRP instance data tree is depicted below:

```
{
  "ietf-network:networks": {
    "nrp-policies": {
      "nrp-policy": [
        {
          "name": "NRP1",
          "nrp-id": "foo:nrp-example1",
          "mode": "nrp-hybrid-plane-partition",
          "resource-reservation": {
            "bw-value": "10000"
          },
          "selector": {
            "ipv6": {
              "ipv6-hbh-eh:": "100"
            }
          },
          "phb-profile:": "High",
          "topology": {
            "igp-congruent": {
              "multi-topology-id": "2"
            },
            "select": {
```

```
              "topology-group": [
                {
                  "group-id": "access-group",
                  "base-topology-ref": {
                    "network-ref": "native-topology"
                  },
                  "link": [
                    {
                      "link-ref": "D1,1-2-1,D2,2-1-1"
                    },
                    {
                      "link-ref": "D2,2-1-1,D1,1-2-1"
                    },
                    {
                      "link-ref": "D1,1-3-1,D3,3-1-1"
                    },
                    {
                      "link-ref": "D3,3-1-1,D1,1-3-1"
                    },
                    {
                      "link-ref": "D2,2-3-1,D3,3-2-1"
                    },
                    {
                      "link-ref": "D3,3-2-1,D2,2-3-1"
                    }
                  ],
                  "link-partition-type": "virtual-sub-interface-partition"
                }
              ]
            }
          }
        }
      ]
    }
  }
}
```

Figure 12: Instance data tree

In addition, an exampe of an NRP that supports the control plane partition mode is shown in the following figure.

```
{
  "ietf-network:networks": {
    "nrp-policies": {
      "nrp-policy": [
        {
          "name": "NRP2",
          "nrp-id": "foo:nrp-example2",
          "mode": "nrp-control-plane-partition",
          "resource-reservation": {
            "bw-value": "10000"
          },
          "phb-profile:": "EF",
          "topology": {
            "filters": {
              "filter": [
                {
                  "filter-ref": "te-topology-filter1"
                }
              ]
            }
          }
        }
      ]
    }
  }
}
```

Appendix C. NRPs YANG Module Tree

Figure 13 shows the full tree diagram of the NRPs YANG model defined in module 'ietf-nrp.yang'.

```
module: ietf-nrp
  augment /nw:networks:
    +--rw nrp-policies
       +--rw nrp-policy* [name]
          +--rw name                        string
          +--rw nrp-id?                     uint32
          +--rw mode?                       identityref
          +--rw resource-reservation
          |  +--rw (max-bw-type)?
          |     +--:(bw-value)
          |     |  +--rw maximum-bandwidth?           uint64
          |     +--:(bw-percentage)
          |        +--rw maximum-bandwidth-percent?
          |                 rt-types:percentage
          +--rw selector
          |  +--rw ipv4
```

```
|  |  +--rw destination-prefix*   inet:ipv4-prefix
|  +--rw ipv6
|  |  +--rw (selector-type)?
|  |     +--:(dedicated)
|  |     |  +--rw ipv6-hbh-eh?          uint32
|  |     +--:(srv6-sid-derived)
|  |     |  +--rw srv6-sid*
|  |     |          inet:ipv6-prefix
|  |     +--:(ipv6-destination-derived)
|  |        +--rw destination-prefix*
|  |                inet:ipv6-prefix
|  +--rw mpls
|  +--rw acl-ref*   nrp-acl-ref
+--rw phb-profile?            string
+--rw topology
   +--rw igp-congruent!
   |  +--rw multi-topology-id?   uint32
   |  +--rw algo-id?             uint32
   |  +--rw sharing?             boolean
   +--rw (topology-type)?
      +--:(selection)
      |  +--rw select
      |     +--rw topology-group* [group-id]
      |        +--rw group-id                string
      |        +--rw base-topology-ref
      |        |  +--rw network-ref?   leafref
      |        +--rw links* [link-ref]
      |        |  +--rw link-ref    leafref
      |        +--rw resource-reservation
      |        |  +--rw (max-bw-type)?
      |        |     +--:(bw-value)
      |        |     |  +--rw maximum-bandwidth?
      |        |     |          uint64
      |        |     +--:(bw-percentage)
      |        |        +--rw maximum-bandwidth-percent?
      |        |                rt-types:percentage
      |        +--rw link-partition-type?
      |        |       identityref
      |        +--rw phb-profile?            string
      +--:(filter)
         +--rw filters
            +--rw filter* [filter-ref]
               +--rw filter-ref
               |       nrp-topo-filter-ref
               +--rw resource-reservation
               |  +--rw (max-bw-type)?
               |     +--:(bw-value)
               |     |  +--rw maximum-bandwidth?
```

```
                              |      |          uint64
                              |      +--:(bw-percentage)
                              |         +--rw maximum-bandwidth-percent?
                              |                   rt-types:percentage
                              +--rw selector
                              |  +--rw ipv4
                              |  |  +--rw destination-prefix*
                              |  |           inet:ipv4-prefix
                              |  +--rw ipv6
                              |  |  +--rw (selector-type)?
                              |  |     +--:(dedicated)
                              |  |     |  +--rw ipv6-hbh-eh?
                              |  |     |           uint32
                              |  |     +--:(srv6-sid-derived)
                              |  |     |  +--rw srv6-sid*
                              |  |     |           inet:ipv6-prefix
                              |  |     +--:(ipv6-destination-derived)
                              |  |        +--rw destination-prefix*
                              |  |                 inet:ipv6-prefix
                              |  +--rw mpls
                              |  +--rw acl-ref*   nrp-acl-ref
                              +--rw phb-profile?            string
  augment /nw:networks/nw:network/nw:network-types:
    +--rw nrp!
  augment /nw:networks/nw:network/nw:node:
    +--ro nrp
       +--ro selector
          +--ro srv6?   srv6-types:srv6-sid
  augment /nw:networks/nw:network/nt:link:
    +--ro nrp
       +--ro link-partition-type?   identityref
       +--ro bandwidth-value?       uint64
       +--ro selector
       |  +--ro srv6?   srv6-types:srv6-sid
       +--ro statistics
          +--ro admin-status?
          |       te-types:te-admin-status
          +--ro oper-status?
          |       te-types:te-oper-status
          +--ro one-way-available-bandwidth?   uint64
          +--ro one-way-utilized-bandwidth?    uint64
          +--ro one-way-min-delay?             uint32
          +--ro one-way-max-delay?             uint32
          +--ro one-way-delay-variation?       uint32
          +--ro one-way-packet-loss?           decimal64
  augment /nw:networks/nw:network/nw:node:
    +--ro nrps* [nrp-id]
       +--ro nrp-id    uint32
```

```
         +--ro nrp
            +--ro selector
               +--ro srv6?   srv6-types:srv6-sid
     augment /nw:networks/nw:network/nt:link:
       +--ro nrps* [nrp-id]
          +--ro nrp-id                uint32
          +--ro link-partition-type?  identityref
          +--ro bandwidth-value?      uint64
          +--ro selector
             +--ro srv6?   srv6-types:srv6-sid
```

Figure 13

Appendix D. NRPs Device YANG Module Tree

Figure 14 shows the full tree diagram of the NRPs device YANG model defined in module 'ietf-nrp-device.yang'.

```
module: ietf-nrp-device
  augment /nw:networks/nrp:nrp-policies/nrp:nrp-policy:
    +--rw interfaces
       +--rw interface* [interface]
          +--rw interface               if:interface-ref
          +--rw resource-reservation
          |  +--rw (max-bw-type)?
          |     +--:(bw-value)
          |     |  +--rw maximum-bandwidth?           uint64
          |     +--:(bw-percentage)
          |        +--rw maximum-bandwidth-percent?
          |                 rt-types:percentage
          +--rw selector
          |  +--rw ipv4
          |  |  +--rw destination-prefix*   inet:ipv4-prefix
          |  +--rw ipv6
          |  |  +--rw (selector-type)?
          |  |     +--:(dedicated)
          |  |     |  +--rw ipv6-hbh-eh?          uint32
          |  |     +--:(srv6-sid-derived)
          |  |     |  +--rw srv6-sid*
          |  |     |           srv6-types:srv6-sid
          |  |     +--:(ipv6-destination-derived)
          |  |        +--rw destination-prefix*
          |  |                 inet:ipv6-prefix
          |  +--rw mpls
          |  +--rw acl-ref*   nrp-acl-ref
          +--rw phb-profile?            string
```

Figure 14

Authors' Addresses

Bo Wu
Huawei Technologies
101 Software Avenue, Yuhua District
Nanjing
Jiangsu, 210012
China
Email: lana.wubo@huawei.com

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park
Bangalore 560066
Karnataka
India
Email: dhruv.ietf@gmail.com

Vishnu Pavan Beeram
Juniper Networks
Email: vbeeram@juniper.net

Tarek Saad
Cisco Systems
Email: tsaad.net@gmail.com

Shaofu Peng
ZTE Corporation
Email: peng.shaofu@zte.com.cn

TEAS Working Group
Internet-Draft
Intended status: Informational
Expires: 29 May 2024

T. Saad
Cisco Systems Inc.
V. Beeram
Juniper Networks
J. Dong
Huawei Technologies
B. Wen
Comcast
D. Ceccarelli
Cisco Systems Inc.
J. Halpern
Ericsson
S. Peng
R. Chen
ZTE Corporation
X. Liu
IBM Corporation
L. Contreras
Telefonica
R. Rokui
Ciena
L. Jalil
Verizon

26 November 2023

Realizing Network Slices in IP/MPLS Networks
draft-ietf-teas-ns-ip-mpls-03

Abstract

Realizing network slices may require the Service Provider to have the ability to partition a physical network into multiple logical networks of varying sizes, structures, and functions so that each slice can be dedicated to specific services or customers. Multiple network slices can be realized on the same network while ensuring slice elasticity in terms of network resource allocation. This document describes a scalable solution to realize network slicing in IP/MPLS networks by supporting multiple services on top of a single physical network by relying on compliant domains and nodes to provide forwarding treatment (scheduling, drop policy, resource usage) on to packets that carry identifiers that indicate the slicing service that is to be applied to the packets.

This Internet-Draft will expire on 29 May 2024.

Table of Contents

1. Introduction

Network slicing allows a Service Provider to create independent and logical networks on top of a shared physical network infrastructure. Such network slices can be offered to customers or used internally by the Service Provider to enhance the delivery of their service offerings. A Service Provider can also use network slicing to structure and organize the elements of its infrastructure. The solution discussed in this document works with any path control technology (such as RSVP-TE, or SR) that can be used by a Service Provider to realize network slicing in IP/MPLS networks.

[I-D.ietf-teas-ietf-network-slices] provides the definition of a network slice for use within the IETF and discusses the general framework for requesting and operating IETF Network Slices, their characteristics, and the necessary system components and interfaces. It also discusses the function of an IETF Network Slice Controller and the requirements on its northbound and southbound interfaces.

This document introduces the notion of a Slice-Flow Aggregate which comprises of one or more IETF network slice traffic streams. It also describes the Network Resource Partition (NRP) and the NRP Policy that can be used to instantiate control and data plane behaviors on select topological elements associated with the NRP that supports a Slice-Flow Aggregate - refer Section 5.1 for further details.

The IETF Network Slice Controller is responsible for the aggregation of multiple IETF network traffic streams into a Slice-Flow Aggregate, and for maintaining the mapping required between them. The mechanisms used by the controller to determine the mapping of one or more IETF network slice to a Slice-Flow Aggregate are outside the scope of this document. The focus of this document is on the mechanisms required at the device level to address the requirements of network slicing in packet networks.

In a Diffserv (DS) domain [RFC2475], packets requiring the same forwarding treatment (scheduling and drop policy) are classified and marked with the respective Class Selector (CS) Codepoint (or the Traffic Class (TC) field for MPLS packets [RFC5462]) at the DS domain ingress nodes. Such packets are said to belong to a Behavior Aggregate (BA) that has a common set of behavioral characteristics or a common set of delivery requirements. At transit nodes, the CS is inspected to determine the specific forwarding treatment to be applied before the packet is forwarded. A similar approach is adopted in this document to realize network slicing. The solution proposed in this document does not mandate Diffserv to be enabled in the network to provide a specific forwarding treatment. If Diffserv is enabled within the network, the Slice-Flow Aggregate traffic can further carry a Diffserv CS to enable differentiation of forwarding treatments for packets within a Slice-Flow Aggregate.

When logical networks associated with an NRP are realized on top of a shared physical network infrastructure, it is important to steer traffic on the specific network resources partition that is allocated for a given Slice-Flow Aggregate. In packet networks, the packets of a specific Slice-Flow Aggregate may be identified by one or more specific fields carried within the packet. An NRP ingress boundary node (where Slice-Flow Aggregate traffic enters the NRP) populates the respective field(s) in packets that are mapped to a Slice-Flow Aggregate in order to allow interior NRP nodes to identify and apply the specific Per NRP Hop Behavior (NRP-PHB) associated with the Slice-Flow Aggregate. The NRP-PHB defines the scheduling treatment and, in some cases, the packet drop probability.

This document covers different modes of NRPs and discusses how each mode can ensure proper placement of Slice-Flow Aggregate paths and respective treatment of Slice-Flow Aggregate traffic.

1.1. Terminology

The reader is expected to be familiar with the terminology specified in [I-D.ietf-teas-ietf-network-slices].

The following terminology is used in the document:

IETF Network Slice:
refer to the definition of 'IETF network slice' in [I-D.ietf-teas-ietf-network-slices].

IETF Network Slice Controller (NSC):
refer to the definition in [I-D.ietf-teas-ietf-network-slices].

Network Resource Partition:
refer to the definition in [I-D.ietf-teas-ietf-network-slices].

Slice-Flow Aggregate:
a collection of packets that are mapped to an NRP and are given the same forwarding treatment; a Slice-Flow Aggregate comprises of one or more IETF network slice traffic streams from one or more connectivity constructs (belonging to one or more IETF network slices); the mapping of one or more IETF network slice streams to a Slice-Flow Aggregate is maintained by the IETF Network Slice Controller. The boundary nodes MAY also maintain a mapping of specific IETF network slice service(s) to a SFA.

Network Resource Partition Policy (NRP):
a policy construct that enables instantiation of mechanisms in support of IETF network slice specific control and data plane behaviors on select topological elements; the enforcement of an NRP Policy results in the creation of an NRP.

NRP Identifier (NRP-ID):
an identifier that is globally unique within an NRP domain and that can be used in the control or management plane to identify the resources associated with the NRP.

NRP Capable Node:
a node that supports one of the NRP modes described in this document.

NRP Incapable Node:
a node that does not support any of the NRP modes described in this document.

Slice-Flow Aggregate Path:

a path that is setup over the NRP that is associated with a specific Slice-Flow Aggregate.

Slice-Flow Aggregate Packet:
a packet that traverses over the NRP that is associated with a specific Slice-Flow Aggregate.

NRP Filtered Topology:
a set of topological elements associated with a Network Resource Partition.

NRP state aware TE (NRP-TE):
a mechanism for TE path selection that takes into account the available network resources associated with a specific NRP.

1.2. Acronyms and Abbreviations

BA: Behavior Aggregate

CS: Class Selector

NRP-PHB: NRP Per Hop Behavior as described in Section 5.1.3

SLA: Service Level Agreements

SLO: Service Level Objectives

SLE: Service Level Expectations

Diffserv: Differentiated Services

MPLS: Multiprotocol Label Switching

LSP: Label Switched Path

RSVP: Resource Reservation Protocol

TE: Traffic Engineering

SR: Segment Routing

VRF: VPN Routing and Forwarding

AC: Attachment Circuit

CE: Customer Edge

PE: Provider Edge

PCEP: Path Computation Element (PCE) Communication Protocol (PCEP)

2. Network Resource Slicing Membership

An NRP that supports a Slice-Flow Aggregate can be instantiated over parts of an IP/MPLS network (e.g., all or specific network resources in the access, aggregation, or core network), and can stretch across multiple domains administered by a provider. The NRP topology may be comprised of dedicated and/or shared network resources (e.g., in terms of processing power, storage, and bandwidth).

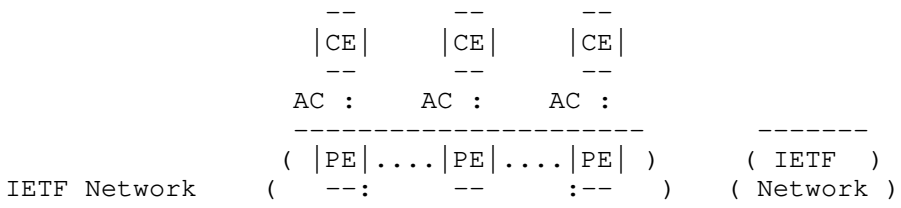
The physical network resources may be fully dedicated to a specific Slice-Flow Aggregate. For example, traffic belonging to a Slice-Flow Aggregate can traverse dedicated network resources without being subjected to contention from traffic of other Slice-Flow Aggregates. Dedicated physical network resource slicing allows for simple partitioning of the physical network resources amongst Slice-Flow Aggregates without the need to distinguish packets traversing the dedicated network resources since only one Slice-Flow Aggregate traffic stream can traverse the dedicated resource at any time.

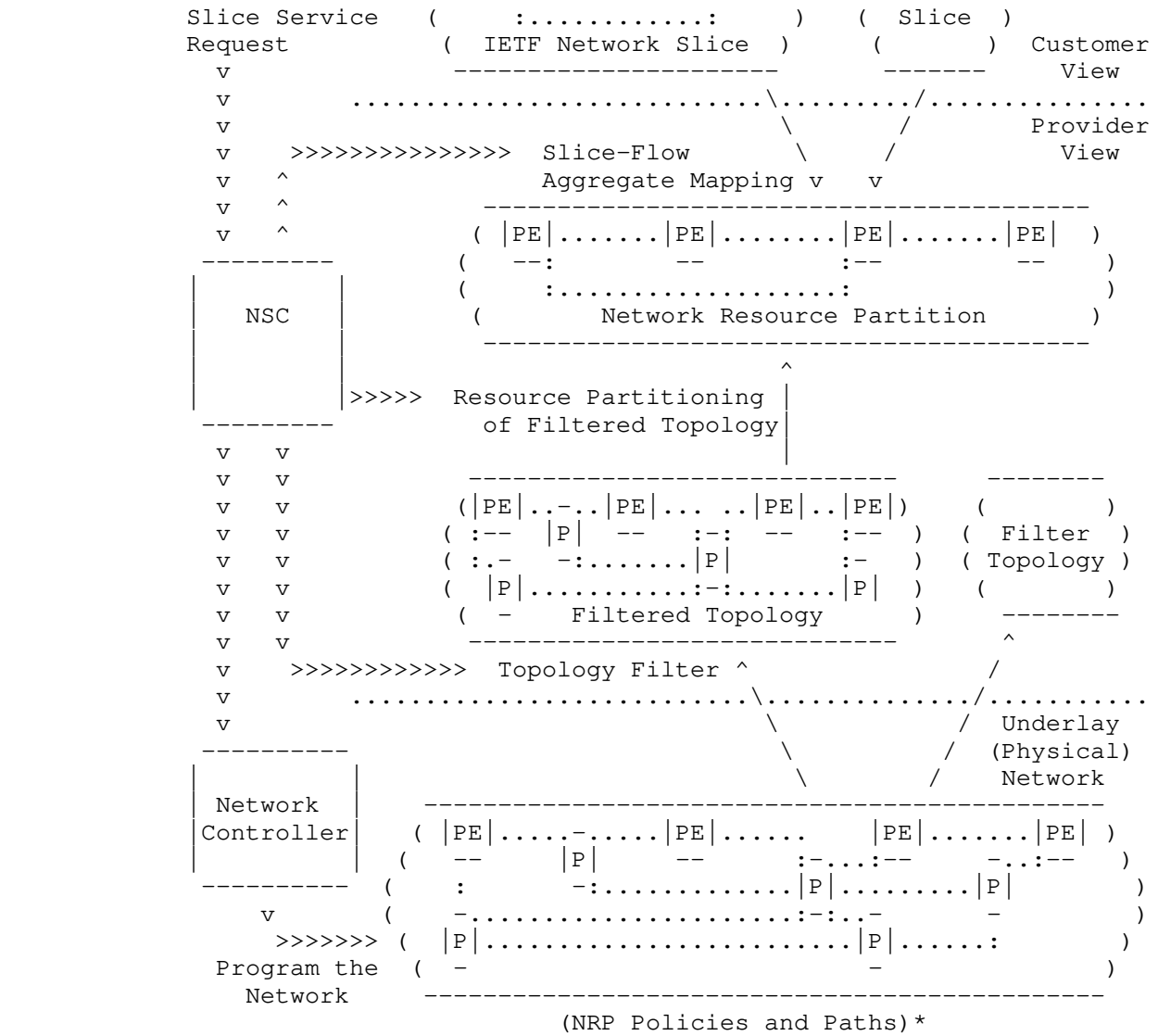
To optimize network utilization, sharing of the physical network resources may be desirable. In such case, the same physical network resource capacity is divided among multiple NRPs that support multiple Slice-Flow Aggregates. The shared physical network resources can be partitioned in the data plane (for example by applying hardware policers and shapers) and/or partitioned in the control plane by providing a logical representation of the physical link that has a subset of the network resources available to it.

3. IETF Network Slice Realization

Figure 1 describes the steps required to realize an IETF network slice service in a provider network using the solution proposed in this document. While Figure 4 of [I-D.ietf-teas-ietf-network-slices] provides an abstract architecture of an IETF Network Slice, this section intends to offer a realization of that architecture specific for IP/MPLS packet networks.

Each of the steps is further elaborated on in a subsequent section.





* : NRP Policy installation and path placement can be centralized or distributed.

Figure 1: IETF network slice realization steps.

3.1. Network Topology Filters

The Physical Network may be filtered into a number of Filter Topologies. Filter actions may include selection of specific nodes and links according to their capabilities and are based on network-wide policies. The resulting topologies can be used to host IETF Network Slices and provide a useful way for the network operator to know that all of the resources they are using to plan a network slice meet specific SLOs. This step can be done offline during planning activity, or could be performed dynamically as new demands arise.

Section 5.1.4 describes how topology filters can be associated with the NRP instantiated by the NRP Policy.

3.2. IETF Network Slice Service Request

The customer requests an IETF Network Slice Service specifying the CE-AC-PE points of attachment, the connectivity matrix, and the SLOs/SLEs as described in [I-D.ietf-teas-ietf-network-slices]. These capabilities are always provided based on a Service Level Agreement (SLA) between the network slice costumer and the provider.

This defines the traffic flows that need to be supported when the slice is realized. Depending on the mechanism and encoding of the Attachment Circuit (AC), the IETF Network Slice Service may also include information that will allow the operator's controllers to configure the PEs to determine what customer traffic is intended for this IETF Network Slice.

IETF Network Slice Service Requests are likely to arrive at various times in the life of the network, and may also be modified.

3.3. Slice-Flow Aggregation

A network may be called upon to support very many IETF Network Slices, and this could present scaling challenges in the operation of the network. In order to overcome this, the IETF Network Slice streams may be aggregated into groups according to similar characteristics.

A Slice-Flow Aggregate is a construct that comprises the traffic flows of one or more IETF Network Slices. The mapping of IETF Network Slices into an Slice-Flow Aggregate is a matter of local operator policy is a function executed by the Controller. The Slice-Flow Aggregate may be preconfigured, created on demand, or modified dynamically.

3.4. Path Placement over NRP Filtered Topology

Depending on the underlying network technology, the paths are selected in the network in order to best deliver the SLOs for the different services carried by the Slice-Flow Aggregate. The path placement function (carried on ingress node or by a controller) is performed on the Filtered Topology that is selected to support the Slice-Flow Aggregate.

Note that this step may indicate the need to increase the capacity of the underlying Filtered Topology or to create a new Filtered Topology.

3.5. NRP Policy

The NRP Policy is a construct that enables the instantiation of control and data plane behaviors on select topological elements in support of the IETF network slice service. The NRP Policy encompasses policy actions (see Section 5.1) that manage the specific resources in the network associated with the NRP.

3.6. NRP Policy Installation

A Controller function programs the physical network with the NRP policies to define specific handling for traffic flows belonging to the Slice-Flow Aggregate. These NRP policies may be consumed on select topological elements in the network and as a result define how routers handle traffic for the Slice-Flow Aggregate associated with the NRP.

For example, the routers that instantiate the NRP Policy can correlate markers that are present in packets that belong to the Slice-Flow Aggregate and apply specific treatments to them.

The way in which the NRP Policy is installed in the routers and the way that the traffic is marked is implementation specific. The NRP Policy instantiation in the network is further described in Section 5.

3.7. Path Instantiation

Depending on the underlying network technology, a Controller function may install the forwarding state specific to the Slice-Flow Aggregate so that traffic is routed along paths derived in the Path Placement step described in Section 3.4. The way in which the paths are instantiated is implementation specific.

3.8. Service Mapping

The edge points can be configured to support the network slice service by mapping the customer traffic to Slice-Flow Aggregates, possibly using information supplied when the IETF network slice service was requested. The edge points may also be instructed to mark the packets so that the network routers will know which policies and routing instructions to apply. The steering of traffic onto Slice-Flow Aggregate paths is further described in Section 6.

4. Network Resource Partition Modes

An NRP Policy can be used to dictate if the network resource partitioning of the shared network resources among multiple Slice-Flow Aggregates can be achieved:

a) in data plane only,

b) in control plane only, or

c) in both control and data planes.

4.1. Data plane Network Resource Partition Mode

The physical network resources can be partitioned on network devices by applying a Per Hop forwarding Behavior (PHB) onto packets that traverse the network devices.

When data plane NRP mode is applied, packets need to be forwarded on the specific NRP that supports the Slice-Flow Aggregate to ensure the proper forwarding treatment dictated in the NRP Policy is applied (refer to Section 5.1 below). In this case, an NRP Selector must be carried in each packet to identify the Slice-Flow Aggregate that it belongs to.

The ingress node of an NRP domain adds an NRP Selector field (if not already present) in each Slice-Flow Aggregate packet. In the data plane NRP mode, the transit nodes within an NRP domain use the NRP Selector to associate packets with a Slice-Flow Aggregate and to determine the Network Resource Partition Per Hop Behavior (NRP-PHB) that is applied to the packet (refer to Section 5.1.3 for further details). The CS MAY be used to apply a Diffserv PHB on to the packet to allow differentiation of traffic treatment within the same Slice-Flow Aggregate.

When data plane only NRP mode is used, routers may rely on a network state independent view of the topology to determine the best paths. In this case, the best path selection dictates the forwarding path of

packets to the destination. The NRP Selector field carried in each packet determines the specific NRP-PHB treatment along the selected path.

4.2. Control Plane Network Resource Partition Mode

Multiple NRPs can be realized over the same set of physical resources. Each NRP is identified by an identifier (NRP-ID) that is globally unique within the NRP domain. The NRP state reservations for each NRP can be maintained on the network element or on a controller.

The network reservation states for a specific partition can be represented in a topology that contains all or a subset of the physical network elements (nodes and links) and reflect the network state reservations in that NRP. The logical network resources that appear in the NRP topology can reflect a part, whole, or in-excess of the physical network resource capacity (e.g., when oversubscription is desirable).

For example, the physical link bandwidth can be divided into fractions, each dedicated to an NRP that supports a Slice-Flow Aggregate. The topology associated with the NRP supporting a Slice-Flow Aggregate can be used by routing protocols, or by the ingress/PCE when computing NRP state aware TE paths.

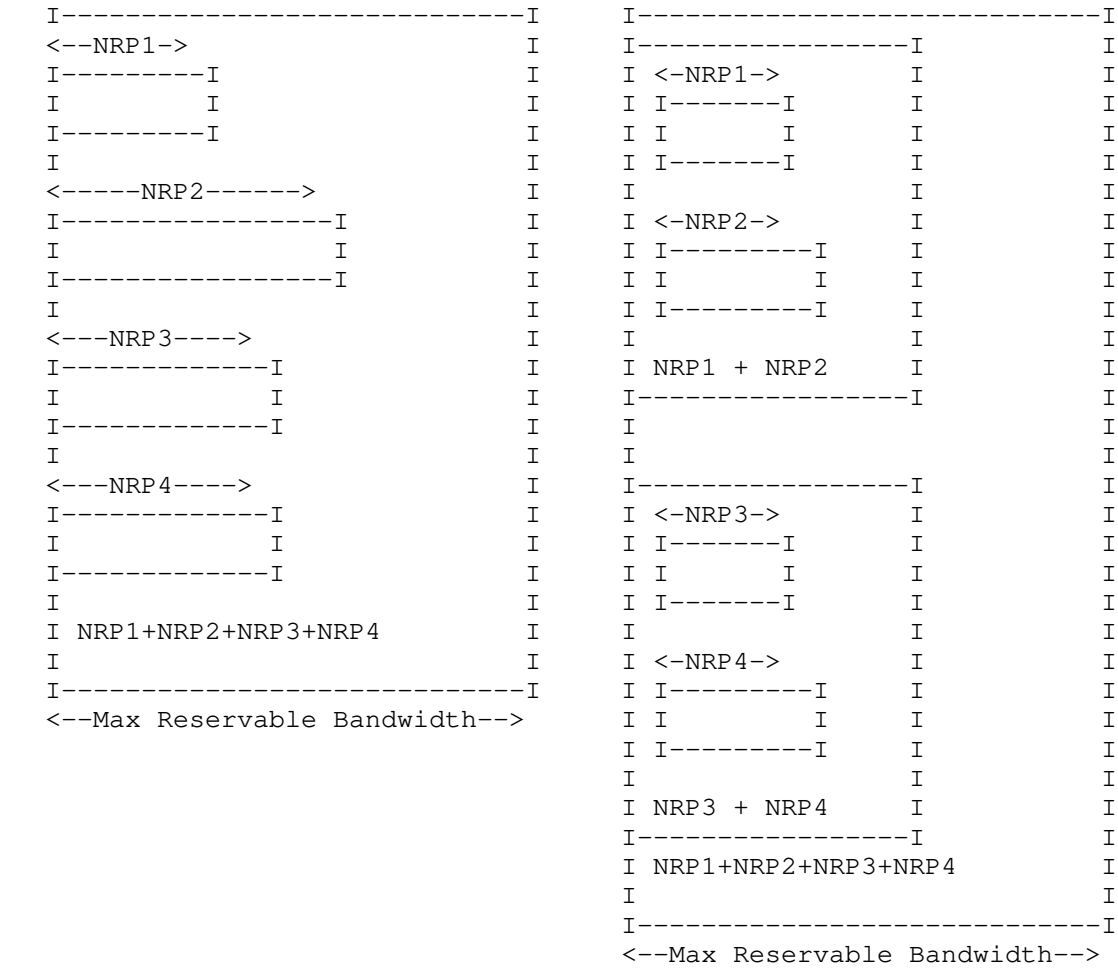
To perform NRP state aware Traffic Engineering (NRP-TE), the resource reservation on each link needs to be NRP aware. The NRP reservations state can be managed locally on the device or off device (e.g. on a controller).

The same physical link may be member of multiple slice policies that instantiate different NRPs. The NRP reservable or utilized bandwidth on such a link is updated (and may be advertised) whenever new paths are placed in the network. The NRP reservation state, in this case, is maintained on each device or off the device on a resource reservation manager that holds reservation states for those links in the network.

Multiple NRPs that support Slice-Flow Aggregates can form a group and share the available network resources allocated to each. In this case, a node can update the reservable bandwidth for each NRP to take into consideration the available bandwidth from other NRPs in the same group.

For illustration purposes, Figure 2 describes bandwidth partitioning or sharing amongst a group of NRPs. In Figure 2a, the NRPs identified by the following NRP-IDs: NRP1, NRP2, NRP3 and NRP4 are

not sharing any bandwidths between each other. In Figure 2b, the NRPs: NRP1 and NRP2 can share the available bandwidth portion allocated to each amongst them. Similarly, NRP3 and NRP4 can share amongst themselves any available bandwidth allocated to them, but they cannot share available bandwidth allocated to NRP1 or NRP2. In both cases, the Max Reservable Bandwidth may exceed the actual physical link resource capacity to allow for over subscription.



(a) No bandwidth sharing between NRPs.

(b) Sharing bandwidth between NRPs of the same group.

Figure 2: Bandwidth isolation/sharing among NRPs.

4.3. Data and Control Plane Network Resource Partition Mode

In order to support strict guarantees for Slice-Flow Aggregates, the network resources can be partitioned in both the control plane and data plane.

The control plane partitioning allows the creation of customized topologies per NRP that each supports a Slice-Flow Aggregate. The ingress routers or a Path Computation Engine (PCE) may use the customized topologies and the NRP state to determine optimal path placement for specific demand flows using NRP-TE.

The data plane partitioning provides isolation for Slice-Flow Aggregate traffic, and protection when resource contention occurs due to bursts of traffic from other Slice-Flow Aggregate traffic that traverses the same shared network resource.

5. Network Resource Partition Instantiation

A network slice can span multiple technologies and multiple administrative domains. Depending on the network slice customer requirements, a network slice can be differentiated from other network slices in terms of data, control, and management planes.

The customer of a network slice service expresses their intent by specifying requirements rather than mechanisms to realize the slice as described in Section 3.2.

The network slice controller is fed with the network slice service intent and realizes it with an appropriate Network Resource Partition Policy (NRP Policy). Multiple IETF network slices are mapped to the same Slice-Flow Aggregate as described in Section 3.3.

The network wide consistent NRP Policy definition is distributed to the devices in the network as shown in Figure 1. The specification of the network slice intent on the northbound interface of the controller and the mechanism used to map the network slice to a Slice-Flow Aggregate are outside the scope of this document and will be addressed in separate documents.

5.1. NRP Policy Definition

The NRP Policy is network-wide construct that is supplied to network devices, and may include rules that control the following:

* Data plane specific policies: This includes the NRP Selector, any firewall rules or flow-spec filters, and QoS profiles associated with the NRP Policy and any classes within it.

* Control plane specific policies: This includes bandwidth reservations, any network resource sharing amongst slice policies, and reservation preference to prioritize reservations of a specific NRP over others.

* Topology membership policies: This defines the topology filter policies that dictate node/link/function membership to a specific NRP.

There is a desire for flexibility in realizing network slices to support the services across networks consisting of implementations from multiple vendors. These networks may also be grouped into disparate domains and deploy various path control technologies and tunnel techniques to carry traffic across the network. It is expected that a standardized data model for NRP Policy will facilitate the instantiation and management of the NRP on the topological elements selected by the NRP Policy topology filter.

It is also possible to distribute the NRP Policy to network devices using several mechanisms, including protocols such as NETCONF or RESTCONF, or exchanging it using a suitable routing protocol that network devices participate in (such as IGP(s) or BGP). The extensions to enable specific protocols to carry an NRP Policy definition will be described in separate documents.

5.1.1. Network Resource Partition - Flow-Aggregate Selector

A router should be able to identify a packet belonging to a Slice-Flow Aggregate before it can apply the associated dataplane forwarding treatment or NRP-PHB. One or more fields within the packet are used as an NRP Selector to do this.

Overloaded forwarding identifier as NRP Selector:

It is possible to assign a different forwarding address (or MPLS forwarding label in case of MPLS network) for each Slice-Flow Aggregate on a specific node in the network. [RFC3031] states in Section 2.1 that: 'Some routers analyze a packet's network layer header not merely to choose the packet's next hop, but also to determine a packet's "precedence" or "class of service"'. Assigning a unique forwarding address (or MPLS forwarding label) to each Slice-Flow Aggregate allows Slice-Flow Aggregate packets destined to a node to be distinguished by the destination address (or MPLS forwarding label) that is carried in the packet.

This approach requires maintaining per Slice-Flow Aggregate state for each destination in the network in both the control and data plane and on each router in the network. For example, consider a

network slicing provider with a network composed of 'N' nodes, each with 'K' adjacencies to its neighbors. Assuming a node can be reached over 'M' different Slice-Flow Aggregates, the node assigns and advertises reachability to 'N' unique forwarding addresses, or MPLS forwarding labels. Similarly, each node assigns a unique forwarding address (or MPLS forwarding label) for each of its 'K' adjacencies to enable strict steering over the adjacency for each slice. The total number of control and data plane states that need to be stored and programmed in a router's forwarding is (N+K)*M states. Hence, as 'N', 'K', and 'M' parameters increase, this approach suffers from scalability challenges in both the control and data planes.

Overloaded service identifier as NRP Selector:

The VPN service label can be overloaded to act as an NRP Selector to allow VPN packets to be mapped to the Slice-Flow Aggregate. In this case, a single VPN service label acting as an NRP Selector needs to be allocated by all Egress PEs of a VPN.

In other cases, a range of VPN service labels can act as an NRP Selector to map VPN traffic to a Slice-Flow Aggregate. An example of such deployment is shown in Figure 3.

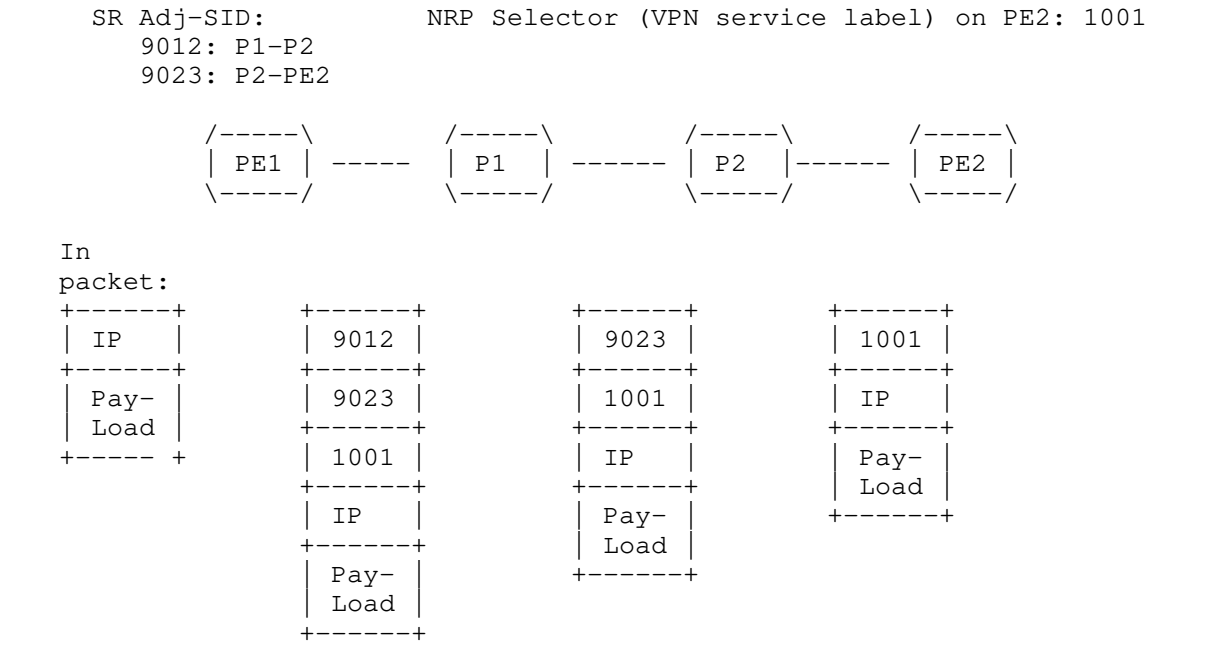


Figure 3: NRP Selector as VPN label at bottom of label stack.

Dedicated identifier as NRP Selector:

An NRP Policy may define a dedicated identifier as NRP Selector that is carried in packets associated with the Slice-Flow Aggregate, independent of the forwarding address or MPLS forwarding label bound to the destination. Routers within the NRP domain can use the forwarding address (or MPLS forwarding label) to determine the forwarding next-hop(s), and use the NRP Selector field in the packet to infer the specific forwarding treatment that needs to be applied on the packet.

The NRP Selector, in this case, can be carried in one of multiple fields in the packet, depending on the dataplane used. For example, in MPLS networks, the NRP Selector can be encoded within an MPLS label that is carried in the packet's MPLS label stack. All packets that belong to the same Slice-Flow Aggregate may carry the same NRP Selector in the MPLS label stack. It is also possible to have multiple NRP Selector's map to the same Slice-Flow Aggregate.

In some cases, the position of the NRP Selector may not be at a fixed position in the MPLS label header. In this case, the NRP Selector label can show up in any position in the MPLS label stack. To enable a transit router to identify the position of the NRP Selector label, a Network Action Indicator (NAI) special purpose label can be used to indicate the presence of a NRP Selector in the MPLS label stack as shown in Figure 4.

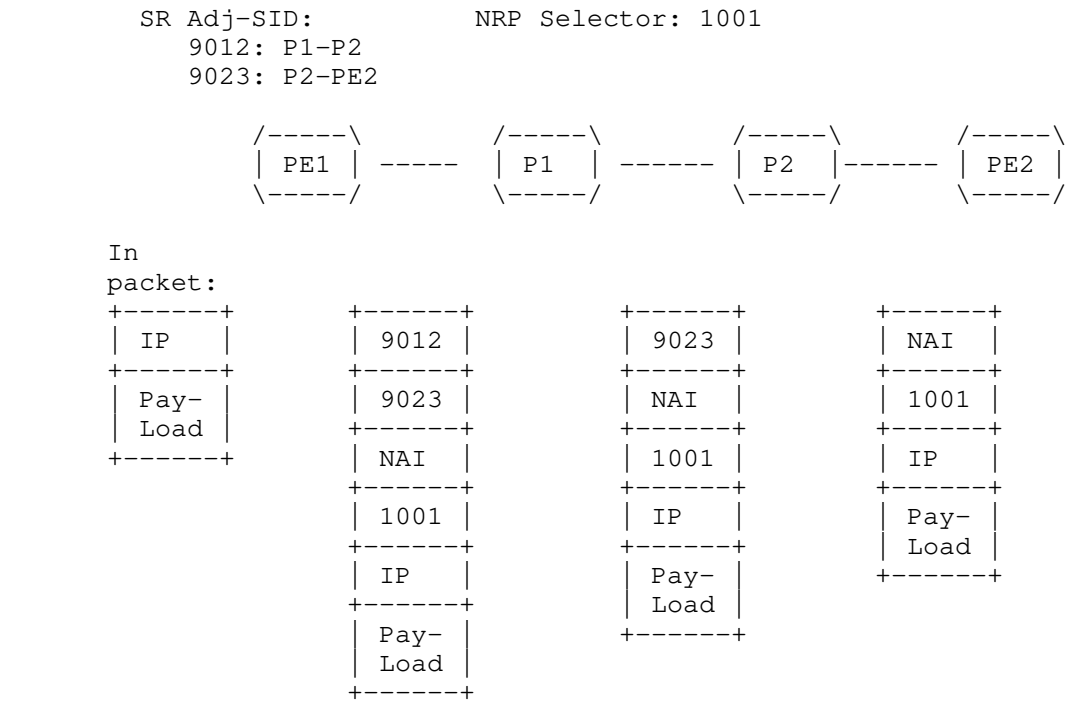


Figure 4: NAI and NRP Selector label in the label stack.

When the slice is realized over an IP dataplane, the NRP Selector can be encoded in the IP header (e.g. as an IPv6 option header).

5.1.2. Network Resource Partition Resource Reservation

Bandwidth and network resource allocation strategies for slice policies are essential to achieve optimal placement of paths within the network while still meeting the target SLOs.

Resource reservation allows for the management of available bandwidth and the prioritization of existing allocations to enable preference-based preemption when contention on a specific network resource arises. Sharing of a network resource's available bandwidth amongst a group of NRPs may also be desirable. For example, a Slice-Flow Aggregate may not be using all of the NRP reservable bandwidth; this allows other NRPs in the same group to use the available bandwidth resources for other Slice-Flow Aggregates.

Congestion on shared network resources may result from sub-optimal placement of paths in different slice policies. When this occurs, preemption of some Slice-Flow Aggregate paths may be desirable to alleviate congestion. A preference-based allocation scheme enables prioritization of Slice-Flow Aggregate paths that can be preempted.

Since network characteristics and its state can change over time, the NRP topology and its network state need to be propagated in the network to enable ingress TE routers or Path Computation Engine (PCEs) to perform accurate path placement based on the current state of the NRP network resources.

5.1.3. Network Resource Partition Per Hop Behavior

The NRP Per Hop Behavior (NRP-PHB) is the externally observable forwarding behavior applied to a specific packet belonging to a Slice-Flow Aggregate. The goal of an NRP-PHB is to provide a specified amount of network resources for traffic belonging to a specific Slice-Flow Aggregate. A single NRP may also support multiple forwarding treatments or services that can be carried over the same logical network.

The Slice-Flow Aggregate traffic may be identified at NRP ingress boundary nodes by carrying a NRP Selector to allow routers to apply a specific forwarding treatment that guarantee the SLA(s).

To support multiple forwarding treatments over the same Slice-Flow Aggregate, a Slice-Flow Aggregate packet may also carry a Diffserv CS to identify the specific Diffserv forwarding treatment to be applied on the traffic belonging to the same NRP.

At transit nodes, the CS field carried inside the packets are used to determine the specific PHB that determines the forwarding and scheduling treatment before packets are forwarded, and in some cases, drop probability for each packet.

5.1.4. Network Resource Partition Topology

A key element of the NRP Policy is a customized topology that may include the full or subset of the physical network topology. The NRP topology could also span multiple administrative domains and/or multiple dataplane technologies.

An NRP topology can overlap or share a subset of links with another NRP topology. A number of topology filtering policies can be defined as part of the NRP Policy to limit the specific topology elements that belong to the NRP. For example, a topology filtering policy can leverage Resource Affinities as defined in [RFC2702] to include or exclude certain links that the NRP is instantiated on in supports of the Slice-Flow Aggregate.

The NRP Policy may also include a reference to a predefined topology (e.g., derived from a Flexible Algorithm Definition (FAD) as defined in [I-D.ietf-lsr-flex-algo], or Multi-Topology ID as defined [RFC4915].

5.2. Network Resource Partition Boundary

A network slice originates at the edge nodes of a network slice provider. Traffic that is steered over the corresponding NRP supporting a Slice-Flow Aggregate may traverse NRP capable as well as NRP incapable interior nodes.

The network slice may encompass one or more domains administered by a provider. For example, an organization's intranet or an ISP. The network provider is responsible for ensuring that adequate network resources are provisioned and/or reserved to support the SLAs offered by the network end-to-end.

5.2.1. Network Resource Partition Edge Nodes

NRP edge nodes sit at the boundary of a network slice provider network and receive traffic that requires steering over network resources specific to a NRP that supports a Slice-Flow Aggregate. These edge nodes are responsible for identifying Slice-Flow Aggregate specific traffic flows by possibly inspecting multiple fields from inbound packets (e.g., implementations may inspect IP traffic's network 5-tuple in the IP and transport protocol headers) to decide on which NRP it can be steered.

Network slice ingress nodes may condition the inbound traffic at network boundaries in accordance with the requirements or rules of each service's SLAs. The requirements and rules for network slice services are set using mechanisms which are outside the scope of this document.

When data plane NRP mode is employed, the NRP ingress nodes are responsible for setting a suitable NRP Selector on packets that belong to the Slice-Flow Aggregate, and optionally the desired Diffserv CS.

5.2.2. Network Resource Partition Interior Nodes

An NRP interior node receives slice traffic and may be able to identify the packets belonging to a specific Slice-Flow Aggregate by inspecting the NRP Selector field carried inside each packet, or by inspecting other fields within the packet that may identify the traffic streams that belong to a specific Slice-Flow Aggregate. For example, when data plane NRP mode is applied, interior nodes can use the NRP Selector carried within the packet to apply the corresponding NRP-PHB forwarding behavior.

5.2.3. Network Resource Partition Incapable Nodes

Packets that belong to a Slice-Flow Aggregate may need to traverse nodes that are NRP incapable. In this case, several options are possible to allow the slice traffic to continue to be forwarded over such devices and be able to resume the NRP forwarding treatment once the traffic reaches devices that are NRP-capable.

When data plane NRP mode is employed, packets carry a NRP Selector to allow slice interior nodes to identify them. To support end-to-end network slicing, the NRP Selector is maintained in the packets as they traverse devices within the network -- including NRP capable and incapable devices.

For example, when the NRP Selector is an MPLS label at the bottom of the MPLS label stack, packets can traverse over devices that are NRP incapable without any further considerations. On the other hand when the NRP Selector label is at the top of the MPLS label stack, packets can be bypassed (or tunneled) over the NRP incapable devices towards the next device that supports NRP as shown in Figure 5.

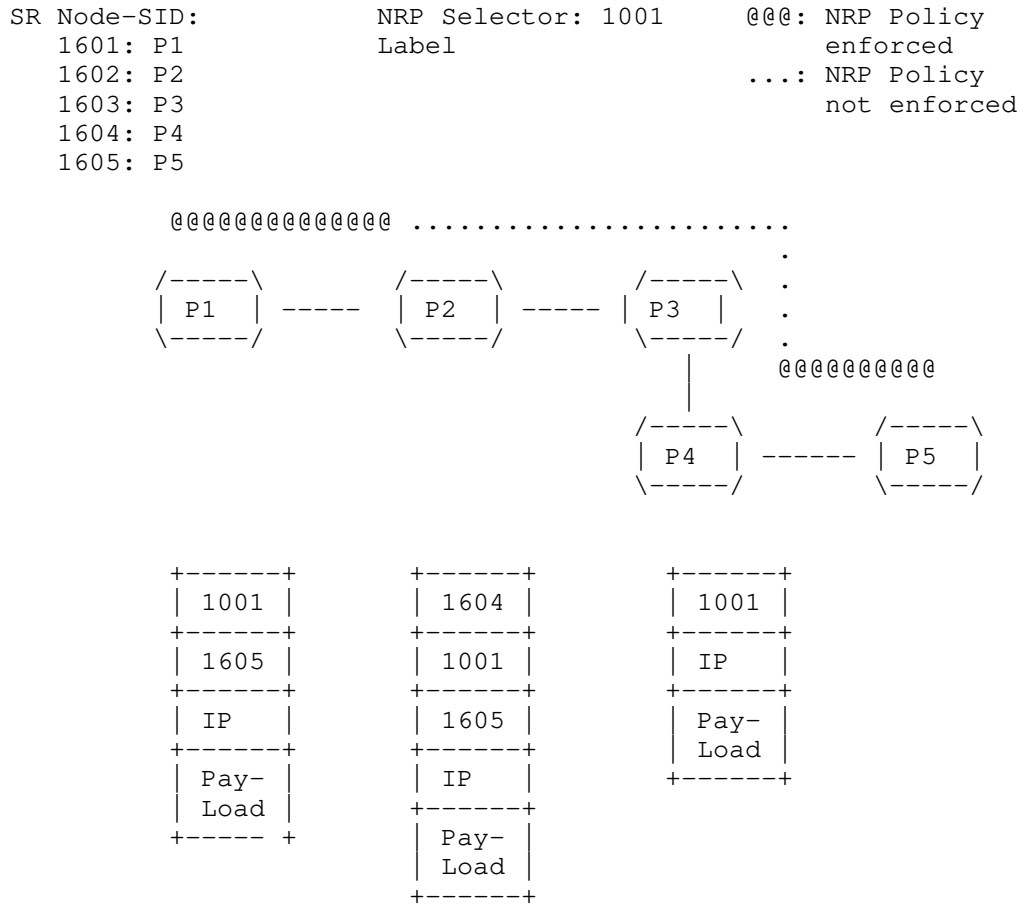


Figure 5: Extending network slice over NRP incapable device(s).

5.2.4. Combining Network Resource Partition Modes

It is possible to employ a combination of the NRP modes that were discussed in Section 4 to realize a network slice. For example, data and control plane NRP modes can be employed in parts of a network, while control plane NRP mode can be employed in the other parts of the network. The path selection, in such case, can take into account the NRP available network resources. The NRP Selector carried within packets allow transit nodes to enforce the corresponding NRP-PHB on the parts of the network that apply the data plane NRP mode. The NRP Selector can be maintained while traffic traverses nodes that do not enforce data plane NRP mode, and so slice PHB enforcement can resume once traffic traverses capable nodes.

6. Mapping Traffic on Slice-Flow Aggregates

The usual techniques to steer traffic onto paths can be applicable when steering traffic over paths established for a specific Slice-Flow Aggregate.

For example, one or more (layer-2 or layer-3) VPN services can be directly mapped to paths established for a Slice-Flow Aggregate. In this case, the per Virtual Routing and Forwarding (VRF) instance traffic that arrives on the Provider Edge (PE) router over external interfaces can be directly mapped to a specific Slice-Flow Aggregate path. External interfaces can be further partitioned (e.g., using VLANs) to allow mapping one or more VLANs to specific Slice-Flow Aggregate paths.

Another option is steer traffic to specific destinations directly over multiple slice policies. This allows traffic arriving on any external interface and targeted to such destinations to be directly steered over the slice paths.

A third option that can also be used is to utilize a data plane firewall filter or classifier to enable matching of several fields in the incoming packets to decide whether the packet belongs to a specific Slice-Flow Aggregate. This option allows for applying a rich set of rules to identify specific packets to be mapped to a Slice-Flow Aggregate. However, it requires data plane network resources to be able to perform the additional checks in hardware.

6.1. Network Slice-Flow Aggregate Relationships

The following describes the generalization relationships between the IETF network slice and different parts of the solution as described in Figure 1.

o A customer may request one or more IETF Network Slices.

o Any given Attachment Circuit (AC) may support the traffic for one or more IETF Network Slices. If there is more than one IETF Network Slice using a single AC, the IETF Network Slice Service request must include enough information to allow the edge nodes to demultiplex the traffic for the different IETF Network Slices.

o By definition, multiple IETF Network Slices may be mapped to a single Slice-Flow Aggregate. However, it is possible for an Slice-Flow Aggregate to contain just a single IETF Network Slice.

o The physical network may be filtered to multiple Filter Topologies. Each such Filtered Topology facilitates planning the placement of paths for the Slice-Flow Aggregate by presenting only the subset of links and nodes that meet specific criteria. Note, however, in absence of any Filtered Topology, Slice-Flow Aggregate are free to operate over the full physical network.

o It is anticipated that there may be very many IETF Network Slices supported by a network operator over a single physical network. A network may support a limited number of Slice-Flow Aggregates, with each of the Slice-Flow Aggregates grouping any number of the IETF Network Slices streams.

7. Path Selection and Instantiation

7.1. Applicability of Path Selection to Slice-Flow Aggregates

In State-dependent TE [I-D.ietf-teas-rfc3272bis], the path selection adapts based on the current state of the network. The state of the network can be based on parameters flooded by the routers as described in [RFC2702]. The link state is advertised with current reservations, thereby reflecting the available bandwidth on each link. Such link reservations may be maintained centrally on a network wide network resource manager, or distributed on devices (as usually done with RSVP-TE). TE extensions exist today to allow IGPs (e.g., [RFC3630] and [RFC5305]), and BGP-LS [RFC7752] to advertise such link state reservations.

When the network resource reservations are maintained for NRPs, the link state can carry per NRP state (e.g., reservable bandwidth). This allows path computation to take into account the specific network resources available for an NRP. In this case, we refer to the process of path placement and path provisioning as NRP aware TE (NRP-TE).

7.2. Applicability of Path Control Technologies to Slice-Flow Aggregates

The NRP modes described in this document are agnostic to the technology used to setup paths that carry Slice-Flow Aggregate traffic. One or more paths connecting the endpoints of the mapped IETF network slices may be selected to steer the corresponding traffic streams over the resources allocated for the NRP that supports a Slice-Flow Aggregate.

The feasible paths can be computed using the NRP topology and network state subject the optimization metrics and constraints.

7.2.1. RSVP-TE Based Slice-Flow Aggregate Paths

RSVP-TE [RFC3209] can be used to signal LSPs over the computed feasible paths in order to carry the Slice-Flow Aggregate traffic. The specific extensions to the RSVP-TE protocol required to enable signaling of NRP aware RSVP-TE LSPs are outside the scope of this document.

7.2.2. SR Based Slice-Flow Aggregate Paths

Segment Routing (SR) [RFC8402] can be used to setup and steer traffic over the computed Slice-Flow Aggregate feasible paths.

The SR architecture defines a number of building blocks that can be leveraged to support the realization of NRPs that support Slice-Flow Aggregates in an SR network.

Such building blocks include:

* SR Policy with or without Flexible Algorithm.

* Steering of services (e.g. VPN) traffic over SR paths

* SR Operation, Administration and Management (OAM) and Performance Management (PM)

SR allows a headend node to steer packets onto specific SR paths using a Segment Routing Policy (SR Policy). The SR policy supports various optimization objectives and constraints and can be used to steer Slice-Flow Aggregate traffic in the SR network.

The SR policy can be instantiated with or without the IGP Flexible Algorithm (Flex-Algorithm) feature. It may be possible to dedicate a single SR Flex-Algorithm to compute and instantiate SR paths for one Slice-Flow Aggregate traffic. In this case, the SR Flex-Algorithm computed paths and Flex-Algorithm SR SIDs are not shared by other Slice-Flow Aggregates traffic. However, to allow for better scale, it may be desirable for multiple Slice-Flow Aggregates traffic to share the same SR Flex-Algorithm computed paths and SIDs.

8. Network Resource Partition Protocol Extensions

Routing protocols may need to be extended to carry additional per NRP link state. For example, [RFC5305], [RFC3630], and [RFC7752] are ISIS, OSPF, and BGP protocol extensions to exchange network link state information to allow ingress TE routers and PCE(s) to do proper path placement in the network. The extensions required to support network slicing may be defined in other documents, and are outside

the scope of this document.

The instantiation of an NRP Policy may need to be automated. Multiple options are possible to facilitate automation of distribution of an NRP Policy to capable devices.

For example, a YANG data model for the NRP Policy may be supported on network devices and controllers. A suitable transport (e.g., NETCONF [RFC6241], RESTCONF [RFC8040], or gRPC) may be used to enable configuration and retrieval of state information for slice policies on network devices. The NRP Policy YANG data model is outside the scope of this document.

9. Outstanding Issues

Note to RFC Editor: Please remove this section prior to publication.

This section records non-blocking issues that were raised during the Working Group Adoption Poll for the document. The below list of issues needs to be fully addressed before progressing the document to publication in IESG.

1. Add new Appendix section with examples for the NRP modes described in Section 4.

2. Elaborate on the SFA packet treatment when no rules to associate the packet to an NRP are defined in the NRP Policy.

3. Clarify how the solution caters to the different IETF Network Slice Service Demarcation Point locations described in Section 4.2 of [I-D.ietf-teas-ietf-network-slices].

4. Clarify the relationship the underlay physical network, the filter topology and the NRP resources.

5. Expand on how isolation between NRPs can be realized depending on the deployed NRP mode.

6. Revise Section 5.2.3 to describe how nodes can discover NRP incapable downstream neighbors.

7. Expand Section 11 on additional security threats introduced with the solution.

8. Expand Section 5.2 on NRP domain boundary and multi-domain aspects.

10. IANA Considerations

This document has no IANA actions.

11. Security Considerations

The main goal of network slicing is to allow for varying treatment of traffic from multiple different network slices that are utilizing a common network infrastructure and to allow for different levels of services to be provided for traffic traversing a given network resource.

A variety of techniques may be used to achieve this, but the end result will be that some packets may be mapped to specific resources and may receive different (e.g., better) service treatment than others. The mapping of network traffic to a specific NRP is indicated primarily by the NRP Selector, and hence an adversary may be able to utilize resources allocated to a specific NRP by injecting packets carrying the same NRP Selector field in their packets.

Such theft-of-service may become a denial-of-service attack when the modified or injected traffic depletes the resources available to forward legitimate traffic belonging to a specific NRP.

The defense against this type of theft and denial-of-service attacks consists of a combination of traffic conditioning at NRP domain boundaries with security and integrity of the network infrastructure within an NRP domain.

12. Acknowledgement

The authors would like to thank Krzysztof Szarkowicz, Swamy SRK, Navaneetha Krishnan, Prabhu Raj Villadathu Karunakaran, and Mohamed Boucadair for their review of this document and for providing valuable feedback on it. The authors would also like to thank Adrian Farrel for detailed discussions that resulted in Section 3.

13. Contributors

The following individuals contributed to this document:

Colby Barth
Juniper Networks
Email: cbarth@juniper.net

Srihari R. Sangli
Juniper Networks
Email: ssangli@juniper.net

Chandra Ramachandran
Juniper Networks
Email: csekar@juniper.net

Adrian Farrel
Old Dog Consulting
United Kingdom
Email: adrian@olddog.co.uk

14. References

14.1. Normative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <https://www.rfc-editor.org/info/rfc3630>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

14.2. Informative References

[I-D.ietf-lsr-flex-algo]
Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", Work in Progress, Internet-Draft, draft-ietf-lsr-flex-algo-26, 17 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-flex-algo-26>.

[I-D.ietf-teas-ietf-network-slices]
Farrel, A., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slices-25, 14 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-25>.

[I-D.ietf-teas-rfc3272bis]
Farrel, A., "Overview and Principles of Internet Traffic Engineering", Work in Progress, Internet-Draft, draft-ietf-teas-rfc3272bis-27, 12 August 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-rfc3272bis-27>.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, DOI 10.17487/RFC2475, December 1998, <https://www.rfc-editor.org/info/rfc2475>.

[RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, DOI 10.17487/RFC2702, September 1999, <https://www.rfc-editor.org/info/rfc2702>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

Authors' Addresses

Tarek Saad
Cisco Systems Inc.
Email: tsaad.net@gmail.com

Vishnu Pavan Beeram
Juniper Networks
Email: vbeeram@juniper.net

Jie Dong
Huawei Technologies
Email: jie.dong@huawei.com

Bin Wen
Comcast
Email: Bin_Wen@cable.comcast.com

Daniele Ceccarelli
Cisco Systems Inc.
Email: dceccare@cisco.com

Joel Halpern
Ericsson
Email: joel.halpern@ericsson.com

Shaofu Peng
ZTE Corporation

Email: peng.shaofu@zte.com.cn

Ran Chen
ZTE Corporation
Email: chen.ran@zte.com.cn

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

Luis M. Contreras
Telefonica
Email: luismiguel.contrerasmurillo@telefonica.com

Reza Rokui
Ciena
Email: rrokui@ciena.com

Luay Jalil
Verizon
Email: luay.jalil@verizon.com