Workgroup: TEAS Working Group
Internet-Draft:
draft-ietf-teas-nrp-scalability-04
Published: 4 March 2024
Intended Status: Informational
Expires: 5 September 2024
Authors: J. Dong
Huawei Technologies
Z. Li
Huawei Technologies
L. Gong
China Mobile
G. Yang
China Telecom
G. Mishra
Verizon Inc.

Scalability Considerations for Network Resource Partition

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 will expire on 5 September 2024.

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

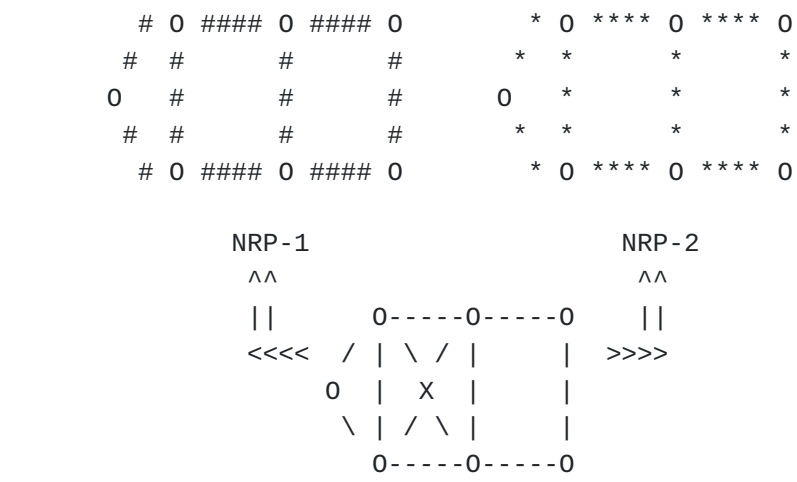
Control plane optimization may be considered in terms of distributed and centralized control planes.

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.

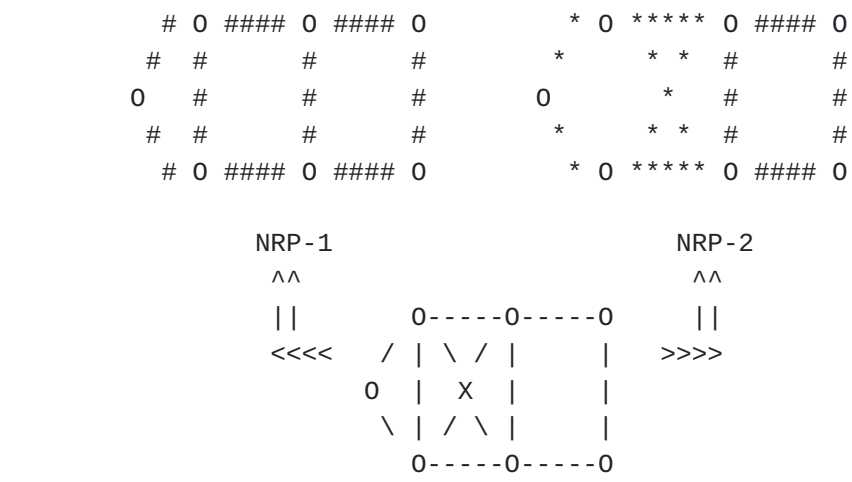


Legend

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.

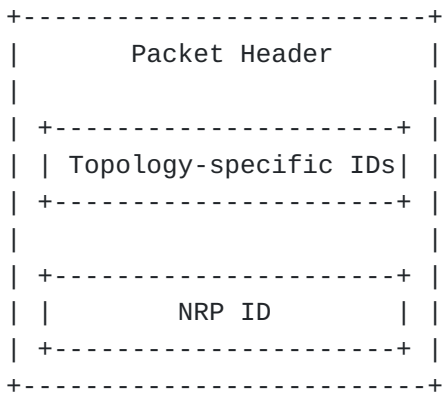


Figure 3. Decoupled Topology and Resource Identifiers in Data Packet

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