TEAS Working Group
Internet-Draft
Intended status: Informational
Expires: January 12, 2022

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

July 11, 2021

Scalability Considerations for Enhanced VPN (VPN+) draft-dong-teas-enhanced-vpn-vtn-scalability-03

Abstract

Enhanced VPN (VPN+) aims to provide enhancements to existing VPN services to support the needs of new applications, particularly including the applications that are associated with 5G services. VPN+ could be used to provide network slicing, and may also be of use in more generic scenarios, such as enterprise services which have demanding requirement. With the requirement for VPN+ services increase, scalability would become an important factor for the deployment of VPN+. This document describes the scalability considerations in the control plane and data plane to enable VPN+ services, some optimization mechanisms are also described.

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 January 12, 2022.

Table of Contents

1. Introduction

Virtual Private Networks (VPNs) have served the industry well as a means of providing different customers with logically isolated connectivity over a common network infrastructure. The common or base network that is used to provide the VPNs is often referred to as the underlay, and the VPN is often called an overlay. The underlay is responsible for establishing the network connectivity and managing network resources to meet the service requirement. The overlay is used to distribute the membership and reachability information of the

customer, and provide logical separation of service delivery between different customers.

Enhanced VPN service (VPN+) [I-D.ietf-teas-enhanced-vpn] is targeted at new applications which require better isolation between customers and/or services, and have more stringent performance requirements than can be provided with existing VPNs. To meet the requirement of VPN+ services, a number of Virtual Transport Networks (VTNs) need to be created, each has a subset of the underlay network topology and a set of network resources allocated from the physical network to meet the requirements of one or a group of VPN+ services. The overlay VPNs together with the corresponding underlay VTN provide the VPN+ service.

Section 6 of [I-D.ietf-teas-enhanced-vpn] provides some general analysis of the scalability of VPN+. This document gives detailed analysis of the scalability considerations when a large number of VPN+ services are provided. Since the scalability of the overlay is not the major bottleneck, this document mainly focuses on the scalability of the underlay VTN.

2. VPN+ Scalability Requirements

As described in [I-D.ietf-teas-enhanced-vpn], VPN+ services may require additional state to be introduced into the network to take advantage of the enhanced functionality. This introduces some scalability considerations to the network. This section gives some analysis of the number of VPN+ services that might be needed in a network.

There are several use cases where VPN+ may be needed, and these determine how many VPN+ will be required in a network. One typical use case of VPN+ is to deliver IETF network slice [I-D.ietf-teas-ietf-network-slices] for applications or services in 5G and other scenarios, thus the number of IETF network slices needed could reflect the number of VPN+ services. With the development and evolution of 5G, it is expected that an increasing number of network slices will be deployed. The number of network slices required depends on how IETF network slices will be used, and the progress of 5G for the vertical industrial services. The potential number of network slices is analyzed by classifying the network slicing deployment into three typical scenarios:

1. Network slices can be used by a network operator internally for different types of services. For example, in a converged multi-service network, different network slices can be created to carry mobile transport service, fixed broadband service and enterprise services respectively, each type of service could be managed by a

separate department or management team. Some service types, such as multicast service may also be deployed in a dedicated network slice. It is also possible that an infrastructure network operator provides network slices to other network operators as a wholesale service. In this scenario, the number of network slices in a network would be relatively small, such as on the order of 10 or so. This could be the typical case in the beginning of the network slice deployment.

2. Network slices can be used to provide isolated and customized virtual networks for customers in different vertical industries. At the early stage of the vertical industrial service deployment, a few top customers in some industries will begin to use network slices to ensure the performance of their business, such as smart grid, manufacturing, public safety, on-line gaming, etc. Considering the number of the vertical industries, and the number of top customers in each industry, the number of network slices may increase to the order of 100.

3. With the evolution of 5G, network slices could be widely used by both vertical industrial customers and enterprise customers which require guaranteed or predictable service performance. The total amount of network slices may increase to the order of 1000 or more. However, it is expected that the number of network slices would still be less than the number of traditional VPN services in the network.

In 3GPP [TS23501], a 5G network slice is identified using Single Network Slice Selection Assistance Information (S-NSSAI), which is a 32-bit identifier comprised of 8-bit Slice/Service Type (SST) and 24-bit Slice Differentiator (SD). This allows the mobile networks (RAN and CN) to provide a large number of network slices. Although it is possible that multiple 5G network slices in RAN and CN are mapped to the same IETF network slice, the number of IETF network slices may still be comparable with the number of 5G network slices. Thus the scalability of IETF network slices needs to be taken into consideration.

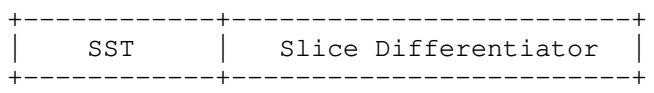

Figure 1. Format of S-NSSAI in 3GPP

VPN+ needs to meet the scalability requirement of network slicing in different scenarios. The increased number of VPN+ services will introduce additional complexity and overhead to both the control

plane and data plane, especially in the aspects related to the underlying VTNs. Although multiple VPN+ services can be mapped to the same VTN as the underlay, there still can be scalability challenges with the increased number of VTNs.

3. VPN+ Scalability Considerations

In this section, the scalability of VTN in the control plane and data plane is analyzed to understand the possible gaps in meeting the scalability requirement of VPN+.

3.1. Control Plane Scalability

As described in [I-D.ietf-teas-enhanced-vpn], the control plane of VPN+ could be based on the hybrid of a centralized controller and the distributed control plane.

3.1.1. Distributed Control Plane

At part of the construction of VPN+ services, it is necessary to create multiple VTNs which provide customized topology and resource attributes. The attributes and state information of each VTN needs to be exchanged in the control plane. The scalability of the distributed control plane for the establishment and maintenance of VTNs needs to be considered in the following aspects:

- The number of control protocol instances maintained on each node
- The number of protocol sessions maintained on each link
- The number of routes advertised by each node
- The amount of attributes associated with each route
- The number of route computation (i.e. SPF computation) executed on each node

As the number of VTNs increases, it is expected that in some of the above aspects, the overhead in the control plane may increase dramatically. For example, the overhead of maintaining separated control protocol instances (e.g. IGP instances) for different VTNs is considered higher than maintaining the information of separated VTNs in the same control protocol instance with appropriate separation, and the overhead of maintaining separate protocol sessions for different VTNs is considered higher than using a shared protocol session for the information exchange of multiple VTNs. To meet the requirement of the increasing number of VTNs, It is

suggested to choose the control plane mechanisms which could improve the scalability while still provide the required functionality.

3.1.2. Centralized Control Plane

Although the SDN approach can reduce the amount of control plane overhead in the distributed control plane, it may transfer some of the scalability concerns from network nodes to the centralized controller, thus the scalability of the controller also needs to be considered.

To provide global optimization for the Traffic Engineered (TE) paths in different VTNs, the controller needs to keep the topology and resource information of all the VTNs up to date. To achieve this, the controller may need to maintain a communication channel with each network node in the network. When there is significant change in the network, or multiple VTNs requires global optimization concurrently, there may be a heavy processing burden at the controller, and a heavy load in the network surrounding the controller for the distribution of the updated network state and the TE paths.

3.2. Data Plane Scalability

To provide different VPN+ services with the required isolation and performance characteristics, it is necessary to allocate different sets of network resources to different VTNs. As the number of VPN+ increases, the number of VTNs will increase accordingly. This requires the underlying network to provide fine-granular network resource partitioning, which means the amount of state about the reserved network resources to be maintained on network nodes will also increase.

In data plane, traffic of different VPN+ services need to be processed separately according to the topology and resource constraints of the associated VTN , thus the information used for VTN identification needs to be carried either directly or implicitly in the data packet. Different approaches of encapsulating the VTN information in data packet can have different scalability implications.

One approach is to reuse some existing fields in the data packet to additionally identify the VTN the packet belongs to. This avoids the cost of introducing new fields in the data packet, while since it introduces additional semantics to an existing field, it requires to change the processing of the existing field in packet forwarding. And when the identifiers which were used to identify a node or link are reused to further identify a VTN, the number of the identifiers

may be increased in proportion to the number of the VTNs, which may cause scalability problem in some networks.

Another alternative approach is to introduce a dedicated field in the packet for VTN identification. This could avoid the impact to the existing fields in the packet. And if this new field carries a global-significant VTN identifier, it could be used together with the existing fields to determine the VTN-specific packet forwarding. The potential issue with this approach is the difficulty in introducing a new field in some types of the data plane.

In addition, the introduction of per VTN packet forwarding has impact on the scalability of the forwarding entries on network nodes, as a network node may need to maintain separate forwarding entries for each VTN it participates in.

3.3. Gap Analysis of Existing Mechanisms

One candidate approach to build VTN is to use VTN specific Segment Routing (either SR-MPLS or SRv6) Identifiers in the data plane [I-D.ietf-spring-sr-for-enhanced-vpn], and define and distribute the associated topology and resource attribute of each VTN based on Multi-topology [RFC4915] [RFC5120] [I-D.ietf-lsr-isis-sr-vtn-mt], Flex-Algo [I-D.ietf-lsr-flex-algo] [I-D.zhu-lsr-isis-sr-vtn-flexalgo] or the combination of these mechanisms in the control plane. This mechanism is suitable for networks with a limited number of VTNs. As the number of VTNs increases, there may be several scalability challenges with this approach:

1. The number of SR SIDs needed will increase in proportion to the number of VTNs in the network, which will bring challenges both to the distribution of SIDs and the related information in the control plane, and to the installation of forwarding entries for VTN-specific SIDs in the data plane.

2. The number of route computation (e.g. SPF computation) will increase in proportion to the number of VTNs in the network, which may introduce significant overhead to the control plane of network nodes.

3. The maximum number of logical topologies supported by OSPF is 128, and the maximum number of Flex-Algo is 128, which may not meet the required number of VTNs in some network scenarios.

4. Possible Scalability Optimizations

4.1. Control Plane Optimizations

For the distributed control plane, several optimizations can be considered to reduce the control plane overhead and improve the scalability.

The first optimization mechanism is to reduce the amount of control plane sessions used for the establishment and maintenance of the VTNs. For multiple VTNs which have the same peering relationship between two adjacent network nodes, it is proposed that one single control protocol session is used for the establishment of multiple VTNs. The information of different VTNs can be exchanged over the same session, with necessary identification information to distinguish the VTNs in the control messages. This could reduce the overhead of maintaining a large number of control protocol sessions for different VTNs, and could also reduce the amount of control plane messages flooded in the network.

The second optimization mechanism is to decompose the attributes of a VTN into different groups, so that different types of VTN attribute can be advertised and processed separately in control plane. There are two basic types of attributes associated with a VTN: the topology attribute and the network resource attribute. In a network, it is possible that multiple VTNs share the same topology, and multiple VTNs may share the same set of network resources on particular network segments. Then it is more efficient if only one copy of the topology attribute is advertised, and multiple VTNs sharing the same topology could refer to this topology information. More importantly, with this approach the result of topology-based route computation could be shared by multiple VTNs, so that the overhead of per-VTN route computation could be reduced . Similarly, information of a subset of network resources reserved on a particular network segment could be advertised once and be referred to by multiple VTNs which share the same set of resources. This methodology could also apply to other attributes of VTN which may be introduced later and can be processed independently.

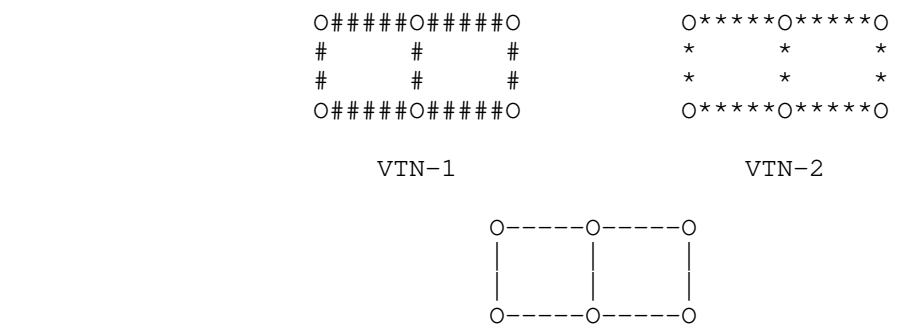

Figure 2. Topology Sharing between VTNs

FIG-2

Figure 2 gives an example of two VTNs which share the same logical topology attribute. As shown in the figure, VTN-1 and VTN-2 have the same topology, while the link resource attributes of each VTN are different. In this case, only one copy of the network topology information needs to be advertised, and the topology-based route computation result can be shared by the two VTNs to generate the corresponding routing and forwarding tables.

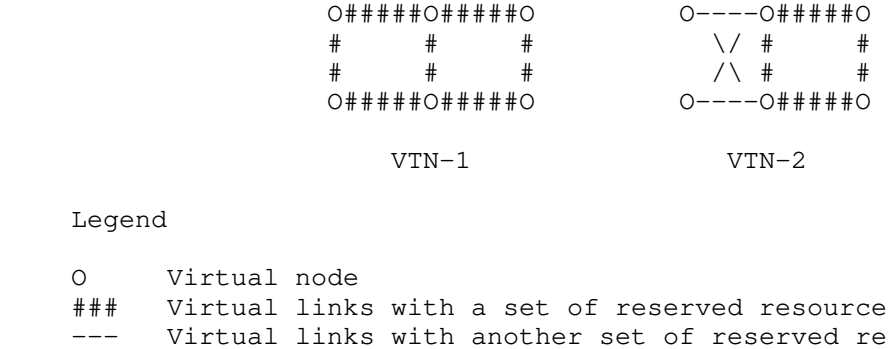

Figure 3. Resource Sharing between VTNs

Figure 3 gives another example of two VTNs which share the same set of network resources on some links. In this case, information about the reserved resource on each link only needs to be advertised once, then both VTN-1 and VTN-2 could refer to the link resource for constraint based path computation.

For the optimization of the centralized control plane, it is suggested that the centralized controller is used as a complementary mechanism to the distributed control plane rather than a replacement, so that the VTN specific path computation burden in control plane could be shared by both the centralized controller and the network nodes, thus the scalability of both systems could be improved.

4.2. Data Plane Optimizations

To support more VPN+ services while keeping the amount of data plane state at a reasonable scale, one possible approach is to classify a set of VPN+ services which have similar service characteristics and performance requirements into a group, and such group of VPN+ services is mapped to one VTN, which is allocated with an aggregated set of network topology and resources to meet the service requirement of the whole group of VPN+. Different groups of VPN+ services need to be mapped to different VTNs with different set of network resources allocated. With appropriate grouping of VPN+ services, a reasonable number of VTNs with network resources reservation and aggregation could still meet the service requirements.

Another optimization in the data plane is to decouple the identifier used for topology-based forwarding and the identifier used for the resource-specific processing introduced by VTN. One possible mechanism is to introduce a dedicated VTN-ID in the packet header to uniquely identify the set of local network resources allocated to a VTN on each network node for the processing and forwarding of the received packet. Then the existing identifier in the packet header used for topology based forwarding is kept unchanged. The benefit is the amount of topology-specific identifiers is in proportion to the number of topologies rather than the number of VTNs, so that its scalability will not be impacted by the increased number of VTN. Since this new VTN-ID field will be used together with the existing fields to determine the VTN-specific packet forwarding, this MAY require network nodes to support a hierarchical forwarding table in the data plane. Figure 4 shows the concept of using different data plane identifiers for topology-based and VTN resource-based packet processing respectively.

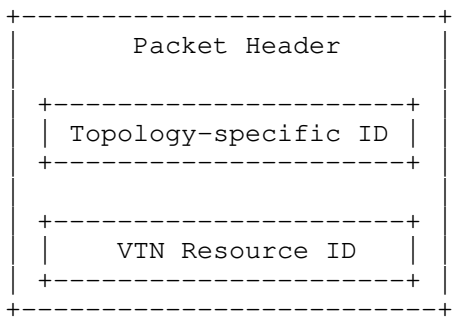

Figure 4. Decoupled Data Plane Identifiers

In an IPv6 [RFC8200] based network, this could be achieved by introducing a dedicated field in either the IPv6 fixed header or the extension headers to carry the VTN identifier for the resource-specific forwarding, while keeping the destination IP address field used for routing towards the destination prefix in the corresponding topology. Note that the VTN-ID needs to be parsed by every node along the path which is capable of VTN-specific forwarding. In an MPLS [RFC3032] based network, this may be achieved by introducing a dedicated MPLS label to identify the VTN, while the existing MPLS labels could be used for topology-based packet forwarding towards the associated destination prefix. This requires that both labels be parsed by each node along the forwarding path of the packet, and the forwarding behavoir depends on the position of the VTN label in the label stack. Another option with the MPLS data plane is to introduce a new MPLS extension header which follows the MPLS label stack to carry the VTN-ID and the associated information. The detailed extensions in IPv6 and MPLS data plane encapsulation are out of the scope of this document.

5. Solution Evolution for Improved Scalability

Based on the analysis in this document, the control plane and data plane for VPN+ needs to evolve to support the increasing number of VPN+ services in the network.

At the first step, by introducing resource-awareness to segment routing SIDs [I-D.ietf-spring-resource-aware-segments], and using Multi-Topology or Flex-Algo as the control plane, it could provide a solution for building a limited number of VTNs in the network to meet the requirement of a relatively small number of VPN+ services in the network. This mechanism is considered as the basic SR VTN.

As the number of required VPN+ services increases, more VTNs may be needed, then the control plane scalability could be improved by

decoupling the topology attribute from other attributes (e.g. resource attribute) of VTN, so that multiple VTNs could share the same topology or resource attribute. This mechanism is considered as the scalable SR VTN. Both the basic and the scalable SR VTN mechanisms are described in [I-D.ietf-spring-sr-for-enhanced-vpn].

If the data plane scalability becomes a concern, dedicated data plane VTN-ID can be introduced to decouple the topology-specific identifiers from the VTN-specific resource identifiers in the data plane, this could help to reduce the number of SR SIDs needed to support a large number of VTNs. This mechanism is considered as the Resource-Independent (RI) VTN.

6. Security Considerations

TBD

7. IANA Considerations

This document makes no request of IANA.

8. Contributors

Zhibo Hu
Email: huzhibo@huawei.com

Hongjie Yang
Email: hongjie.yang@huawei.com

9. Acknowledgments

The authors would like to thank Adrian Farrel for the review and discussion of this document.

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 Enhanced Virtual Private Network (VPN+) Services", draft-ietf-teas-enhanced-vpn-07 (work in progress), February 2021.

10.2. Informative References

[I-D.ietf-lsr-flex-algo]
Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", draft-ietf-lsr-flex-algo-15 (work in progress), April 2021.

[I-D.ietf-lsr-isis-sr-vtn-mt]
Xie, C., Ma, C., Dong, J., and Z. Li, "Using IS-IS Multi-Topology (MT) for Segment Routing based Virtual Transport Network", draft-ietf-lsr-isis-sr-vtn-mt-00 (work in progress), March 2021.

[I-D.ietf-spring-resource-aware-segments]
Dong, J., Bryant, S., Miyasaka, T., Zhu, Y., Qin, F., Li, Z., and F. Clad, "Introducing Resource Awareness to SR Segments", draft-ietf-spring-resource-aware-segments-02 (work in progress), February 2021.

[I-D.ietf-spring-sr-for-enhanced-vpn]
Dong, J., Bryant, S., Miyasaka, T., Zhu, Y., Qin, F., Li, Z., and F. Clad, "Segment Routing based Virtual Transport Network (VTN) for Enhanced VPN", draft-ietf-spring-sr-for-enhanced-vpn-00 (work in progress), February 2021.

[I-D.ietf-teas-ietf-network-slices]
Farrel, A., Gray, E., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "Framework for IETF Network Slices", draft-ietf-teas-ietf-network-slices-00 (work in progress), April 2021.

[I-D.zhu-lsr-isis-sr-vtn-flexalgo]
Zhu, Y., Dong, J., and Z. Hu, "Using Flex-Algo for Segment Routing based VTN", draft-zhu-lsr-isis-sr-vtn-flexalgo-02 (work in progress), February 2021.

[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>.

[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>.

[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>.

[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

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

James N Guichard
Futurewei Technologies
2330 Central Express Way
Santa Clara
USA

Email: james.n.guichard@futurewei.com

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Fengwei Qin
China Mobile
No. 32 Xuanwumenxi Ave., Xicheng District
Beijing
China

Email: qinfengwei@chinamobile.com