

NV03 WG
Internet-Draft
Intended status: Standards Track
Expires: December 7, 2019

T. Ao
ZTE Corporation
G. Mirsky
ZTE Corp.
Y. Fan
China Telecom
June 5, 2019

Architecture for Overlay Virtual Sub-Network Interconnection
draft-ao-nv03-overlay-subnet-architecture-02

Abstract

An virtualized overlay network may be divided into several subnets for the reasons of geographical location, management, or using different technologies being used. For example, different customer have their own preference. But all these subnets need to work together to provide an end-to-end connection in a virtual network, An extended architecture of the NV03 and propose a new component to provide the connection fucntion are introduced in this document.

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 December 7, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Background	2
2.	Conventions used in this document	3
2.1.	Terminology	3
2.2.	Requirements Language	4
3.	Architecture for subnet	4
3.1.	Stitching NVE	5
4.	IANA Considerations	7
5.	References	7
5.1.	Normative References	7
5.2.	Informational References	8
	Authors' Addresses	8

[1.](#) Background

Network virtualization using Overlays over Layer 3 (NV03) is a technology that is used to address issues that arise in building large, multi-tenant data centers that make extensive use of server virtualization.

As described [[I-D.defoy-coms-subnet-interconnection](#)] and [[I-D.homma-coms-slice-gateway](#)], for network slicing in 5G, there are number of reasons to compose an end-to-end network slice instance by using subnets and stitching operations, thus enabling hierarchical/recursive management of slices. These subnets are the part of the network slice instance, but not isolate from network slice. An overlay virtual network may also be constructed of several subnets. In such scenario, subnets interconnection to a virtual network is required. We also can consider such interconnection as stitching. With such stitching, several subnets can provide a end to end connection in an overlay virtual network.

Moreover, with the progress in NV03 WG, some of the data plane encapsulations have been put forward, some are outstanding dataplane for an overlay network, such as VxLAN-GPE [[I-D.ietf-nvo3-vxlan-gpe](#)], GENEVE [[I-D.ietf-nvo3-geneve](#)] and GUE [[I-D.ietf-nvo3-encap](#)], etc. The consideration about these overlay encapsulations has been analyzed in [[I-D.ietf-nvo3-encap](#)]. The fact is that each of them has its customers, and furthermore, some of them have been already deployed in the network. So that a problem arises: for a virtual network, all the hosts that connect to the same VN and want to

communicate with each other are required to have the same data plane encapsulation. This problem limits the network scalability and capacity. Especially, when the NVE is located on the vSwitch, the encapsulation method on the NVE is not predictable. Allowing as many types of accession as possible is more attractive for a virtualized overlay network. So in such a scenario, the NVEs using same technology can be connected to the same subnet.

To improve the scalability and capacity of the virtualized overlay network and to satisfy the subnet interconnection requirement in network slicing, we propose a subnet interconnection architecture in NV03, and provide a stitching gateway between the subnets in a virtual network in this document. With such architecture, the subnets in an overlay virtual network with different technologies and with separate management can be interconnected. The gateway that provides the connection between subnets is referred to as Stitching Gateway. In particular, the Stitching Gateway generally would be located in a certain kind of NVE, such NVE is called as S-NVE in the following description.

2. Conventions used in this document

2.1. Terminology

NV03: Network Virtualization using Overlay over Layer 3

NVA: Network Virtualization Authority

TS: Tenant System

VxLAN-GPE: Virtual extension LAN with Generic Protocol Extension

GENEVE: Generic Network Virtualization Encapsulation

GUE: Generic UDP Encapsulation

S-GW: Stitching Gateway. A gateway that does the stitching for separate subnet to enable them to communicate with each other.

S-NVE: A NVE that complete the stitching functionn as a Stitching Gateway for subnets in an overlay virtual network.

Subnet: An Overlay Virtual Network is combined with several Subnets which may use different technology or different management.

Network slice in this document is used as shortened version of network slice instance.

2.2. 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 [BCP 14](#) [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Architecture for subnet

As Generic Reference Model described in [RFC7364], it's a DC reference model for network virtualization overlays where NVEs provide a logical interconnect between Tenant Systems that belong to a specific VN. But if we need to support an overlay virtual network, an extended reference model is shown as follows in the NV03 architecture.

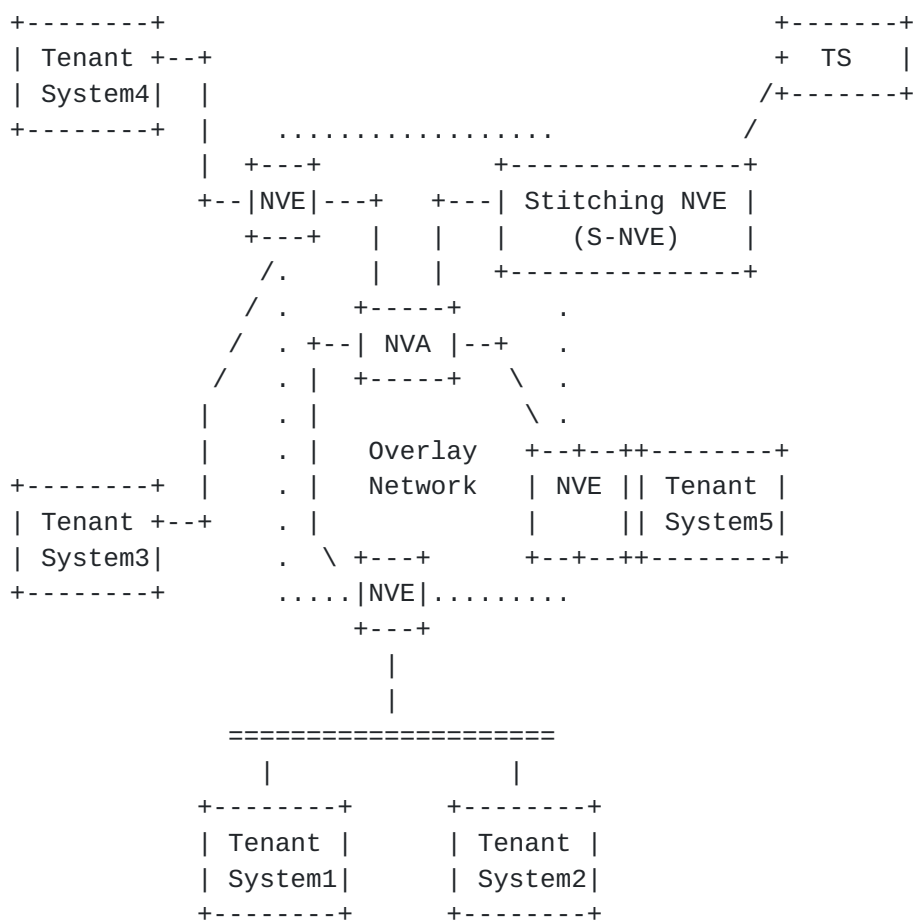


Figure 1 Reference Model supporting subnet

Figure 1: Reference Model supporting subnet

In the above figure, a specific Stitching Gateway(S-Gateway) component is introduced. Generally, the gateway is located on a NVE, so we call it as S-NVE. For the TSs in the same virtual network, if the NVEs which they are connected to are using different overlay technology, want to communicate with each other, Stitching NVE(S-NVE) should take over responsibility as a gateway to provide a "bridge" for the communication. Or for the TSs connecting to different subnets, but belonging to a virtual network, need to communicate with each other through S-NVE.

The difference between NVE and S-NVE is that NVE is the connection between different VN, while S-NVE is the connection between subnet. From the view of NVE, S-NVE is an intermediate relay node. For the two Tenant Systems belong to different subnet have to communicate through a S-NVE, even though they belong to a same virtual network.

That is, when different NVEs want to set up tunnel, if they can't connect each other directly because they are on the different subnets, they can set up a tunnel with S-NVE separately, so that the S-NVE connects the two tunnels as a stitcher. There could be more than one S-NVE in a virtual network.

3.1. Stitching NVE

Stitching NVE(S-NVE) is a certain kind of NVE that maybe appointed by NVA or by the manager. As an essential component in the NVO supporting subnet, the requirements for a Stitching NVE is :

1. Provide the connection for the two subnet of a virtual network.
2. Provides identification/ classification of customer and service traffic, performs the mapping of the two tunnels.
3. Support at least two kinds of encapsulations, translation between technologies/encapsulations when it is stitching two subnets with two different technology .
4. Fault and performance monitoring for underlay and overlay networks.

Regarding the [[RFC8014](#)], the Stitching NVE has a reference model as showed in Figure 2.

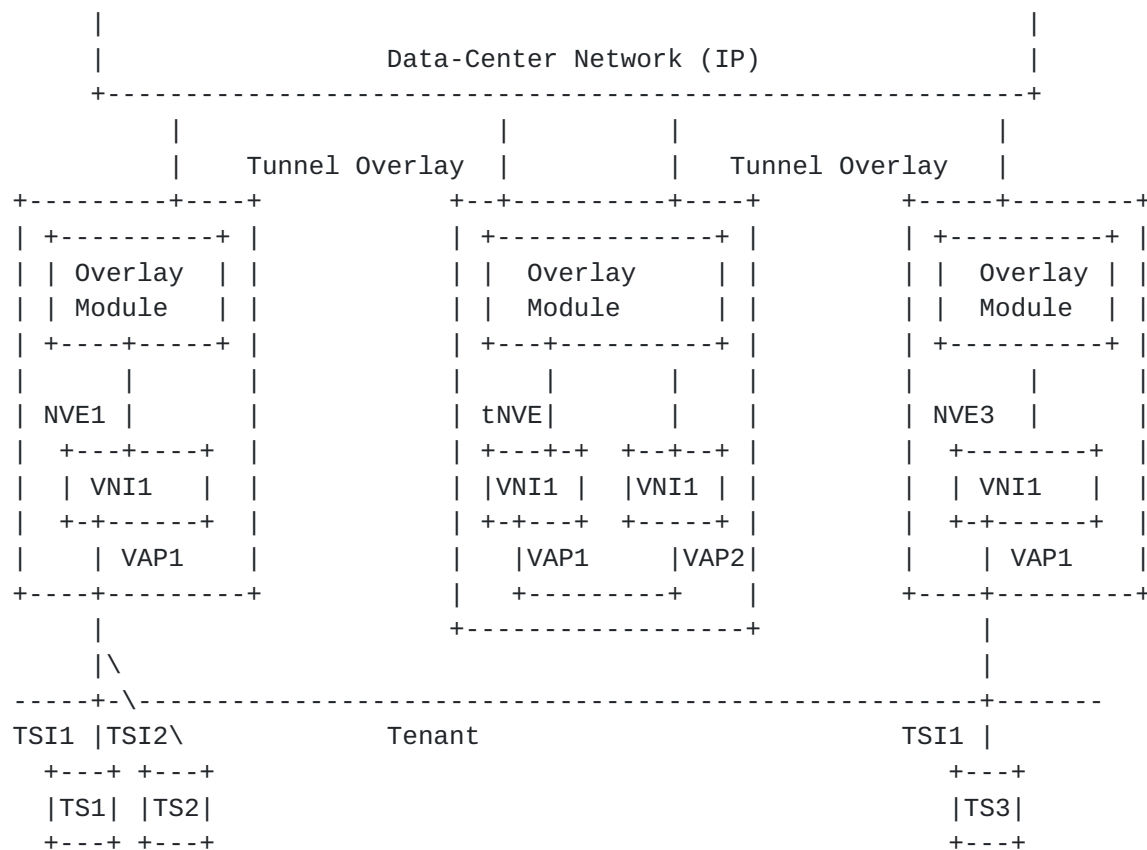


Figure 2 S-NVE Reference Model

Figure 2: S-NVE reference model

S-NVE is a key component of the connection between NVE1 and NVE2. It can be a dedicated device and be a NVE that also provide the overlay network for the TSs. When the NVE takes the role of stitching different subnets for different TSs, it will not forward the traffic to TS, but to another VAP that supports the encapsulation the destination NVE owned.

Take the Figure 2 as an example to illustrate how does S-NVE work. NVE1 belongs to subnet1 and only support VxLAN-GPE, and NVE2 belongs to subnet2 and only support GENEVE. For the two communicating TSs: TS1 needs to send packets to TS3, and TS3 also needs to reply to TS1. They are in the same VNID1, but the NVE they are connected to is using a different encapsulation, and the they belongs to two different subnet. So if the two TSes want to communicate with each other, packets have to transfer at S-NVE first. For NVE1, it has no sense that TS3 is connecting to NVE3, instead of assuming that TS3 is connecting to S-NVE. In the same way, for NVE3, it has no sense that TS1 is connecting to NVE1, instead of assuming that TS1 is connecting to S-NVE. So because of the existence of the tNVE, no matter TS1/TS3 or NVE1/NVE3, they never perceive that they are in the different data

plane. NVE1 getting the packets from TS1 encapsulates them in Vxlan-GPE and then send the packets to S-NVE. The S-NVE receives the packets from the Vxlan-GPE tunnel and then de-encapsulate the vxLAN-GPE to VAP1. Next, the S-NVE forwards packets to the Overlay Module from VAP2 to have another encapsulation GENEVE on the packets. At last S-NVE forwards the packet in the GENEVE tunnel to NVE3.

From the above, S-NVE is like a stitcher between TS1 and TS2. And owing to S-NVE, even though NVE1 and NVE2 which TS1 and TS2 connecting belong to separate subnets and separately have different encapsulation, as long as they are in the same virtual network, they would communicate each other as a Larger L2 network and no need to know that they are in different subnets or using different technology.

4. IANA Considerations

TBD.

5. References

5.1. Normative References

- [I-D.ietf-intarea-gue]
Herbert, T., Yong, L., and O. Zia, "Generic UDP Encapsulation", [draft-ietf-intarea-gue-07](#) (work in progress), March 2019.
- [I-D.ietf-nvo3-geneve]
Gross, J., Ganga, I., and T. Sridhar, "Geneve: Generic Network Virtualization Encapsulation", [draft-ietf-nvo3-geneve-13](#) (work in progress), March 2019.
- [I-D.ietf-nvo3-vxlan-gpe]
Maino, F., Kreeger, L., and U. Elzur, "Generic Protocol Extension for VXLAN", [draft-ietf-nvo3-vxlan-gpe-07](#) (work in progress), April 2019.
- [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>>.
- [RFC7365] Lasserre, M., Balus, F., Morin, T., Bitar, N., and Y. Rekhter, "Framework for Data Center (DC) Network Virtualization", [RFC 7365](#), DOI 10.17487/RFC7365, October 2014, <<https://www.rfc-editor.org/info/rfc7365>>.

- [RFC8014] Black, D., Hudson, J., Kreeger, L., Lasserre, M., and T. Narten, "An Architecture for Data-Center Network Virtualization over Layer 3 (NV03)", [RFC 8014](#), DOI 10.17487/RFC8014, December 2016, <<https://www.rfc-editor.org/info/rfc8014>>.
- [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>>.

5.2. Informational References

- [I-D.ao-nvo3-multi-encap-interconnect]
Ao, T., Mirsky, G., and F. Yongbing, "Multi-encapsulation interconnection for Overlay Virtual Network", [draft-ao-nvo3-multi-encap-interconnect-00](#) (work in progress), March 2018.
- [I-D.defoy-coms-subnet-interconnection]
Foy, X., Rahman, A., Galis, A., kiran.makhijani@huawei.com, k., Qiang, L., Homma, S., and P. Martinez-Julia, "Interconnecting (or Stitching) Network Slice Subnets", [draft-defoy-coms-subnet-interconnection-03](#) (work in progress), March 2018.
- [I-D.homma-coms-slice-gateway]
Homma, S., Foy, X., and A. Galis, "Gateway Function for Network Slicing", [draft-homma-coms-slice-gateway-01](#) (work in progress), March 2018.
- [I-D.ietf-nvo3-encap]
Boutros, S., "NV03 Encapsulation Considerations", [draft-ietf-nvo3-encap-02](#) (work in progress), September 2018.
- [RFC7364] Narten, T., Ed., Gray, E., Ed., Black, D., Fang, L., Kreeger, L., and M. Napierala, "Problem Statement: Overlays for Network Virtualization", [RFC 7364](#), DOI 10.17487/RFC7364, October 2014, <<https://www.rfc-editor.org/info/rfc7364>>.

Authors' Addresses

Ting Ao
ZTE Corporation
No.889, BiBo Road
Shanghai 201203
China

Phone: +86 17721209283
Email: 18555817@qq.com

Greg Mirsky
ZTE Corp.
1900 McCarthy Blvd. #205
Milpitas, CA 95035
USA

Email: gregimirsky@gmail.com

Yongbin
China Telecom
No.109, Zhongshan Avenue
Guangzhou 510630
China

Email: fanyb@gsta.com

