

NVO3
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2016

Z. Gu
Independent Researcher
B. Khasnabish
ZTE (TX) Inc.
T. Ao
Fangwei. Hu
ZTE Corp
October 19, 2015

Virtual Network Transport Protocol (VNTP)
draft-gu-nvo3-vntp-03

Abstract

This document describes the overlay Virtual Network Transport Protocol (VNTP), which defines the interactions between NVE and NVA/NVE and the relevant message to support virtual network implementation.

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 <http://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 April 21, 2016.

Copyright Notice

Copyright (c) 2015 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 (<http://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. Introduction	2
2. Conventions Used in This Document	2
3. VNTP Overview	3
4. VNTP Message Format	3
4.1. VNTP Header format	4
4.2. VNTP Data Format	6
5. The Operations of NVE	9
6. The operations of NVA	9
7. Interaction with TS/Hypervisor-NVE protocol	10
8. Security Considerations	10
9. IANA/IEEE Considerations	10
10. References	10
10.1. Normative references	10
10.2. Informative References	11
Authors' Addresses	11

1. Introduction

[RFC7364] and [RFC7365] describe the need and some characteristics of the interaction between NVE and NVA. [draft-ietf-nvo3-arch-03] has a more detailed architectural description about NVE-NVA protocol. And [draft-ietf-nvo3-nve-nva-cp-req-03] discusses the detail requirements of NVE-NVA protocol.

This draft defines a NVE-NVA protocol, Virtual Network Transport Protocol (VNTP). It belongs to the second model mentioned in [draft-ietf-nvo3-arch-03], e.g. NVE interacts with NVA directly. It defines the interactions between NVE and NVA/NVE and the relevant message formats to support virtual network implementation and fulfill the requirements described in the related documents mentioned above.

VNTP can be based on a broad transport mechanism such as TCP or UDP, or even IP. A new TCP/UDP port or protocol number allocation is needed if the transport mechanism is decided by NVO3 WG.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. VNTP Overview

VNTP based on some basic assumptions and the main points include:

- 1), the first-hand mapping information provided by NVE, either configured by administrator or automatically created. Architecturally, VNTP also support other mapping resources such as downloaded from NVA.
- 2), NVE registers to NVA per VN and NVA may store two lists of NVE, the first one is all NVEs in a VN and another one is about all the VNs NVE resides.
- 3), when mapping change occurs in NVE, the NVE send update message to NVA to initiate the synchronization procedures and NVA then forward the update message to all other NVEs in the same VN. Optionally, NVA can store all the update information for latter use.
- 4), when a NVE register to a VN and some update messages received by NVA, the NVA may use the messages stored or request the related NVEs to send the update again to synchronize.
- 5), if NVA obtains the mapping information from other resources different from NVE, for example configured by administrator or from VM Orchestration, it sends the mapping/update information to all NVEs in the same VN.

The VNTP procedures can be simplified to implement by point-to-point communication between NVE and NVA. So the NVE-NVA interaction can be based on a broad transport mechanism such as TCP, UDP or even IP. A new TCP/UDP port or protocol number allocation is needed if the transport mechanism is decided.

4. VNTP Message Format

Figure 1 shows the VNTP message format. VNTP message format definition is based on some transport mechanism, such as TCP or UDP transport protocol, or even based on IP, and further using its data/payload field.

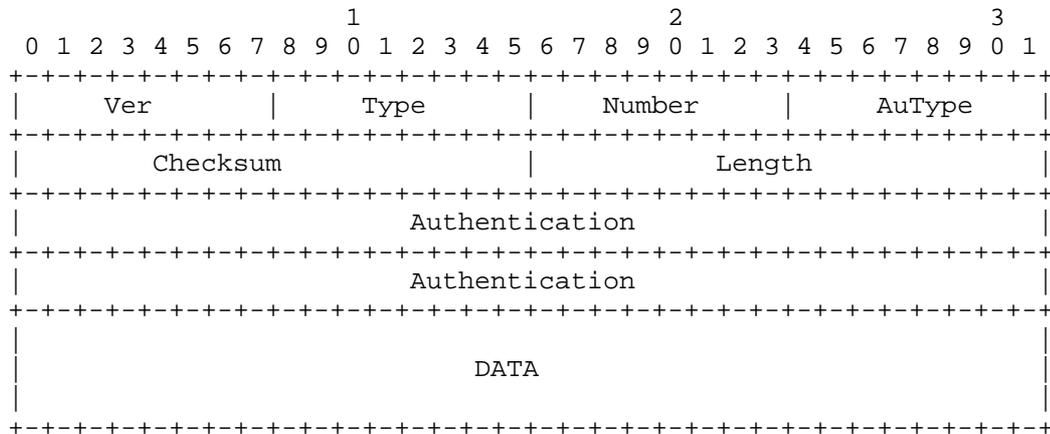


Figure 1 VNTP message format

Figure 1

4.1. VNTP Header format

The following are the Header fields definition.

Ver (8bit): for VNTP version.

Type (8bit): for VNTP Type Command or result/response definition.

Number (8bit): item/entry number of data field.

AuType and Authentication(length TBD): for authentication type and packet authentication.(Refer to RFC2328, especially section A and D; and further to RFC5709 for authentication update discussion.)

Checksum (16bit): checksum of the whole VNTP packet except authentication field.

Length (16bit): total packet octets including header.

Figure 2 shows detail Type definition.

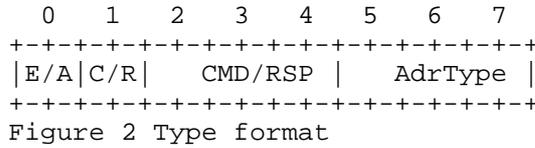


Figure 2

E/A (1bit): set to 1, NVE->NVA;

set to 0, NVA->NVE.

C/R (1bit): 1, CMD/RSP represents Command;

0, CMD/RSP represents Response/Result.

CMD/RSP: A command from NVE or NVA, or a response to the command

Detailed definition for the Command from NVE (C/R = 1, E/A = 1) .

CMD Description

000 NVE registration: The NVE registers to VNs.

001 NVE deregistration: The NVE de-registers from VNs.

010 NVE Update: NVE's mapping information has been update.

011-111: Reserved for future use

Detailed definition for the Command from NVA (C/R = 1, E/A = 0)

CMD Description

000 Request for NVE Mapping information

001 Nullify NVE Mapping information/NVE deregistration

010 (NVE registered) Update NVE mapping information

011-111: Reserved for future use

Detailed definition for NVA/NVE Response/Result (C/R = 0, E/A = 0/1)

RSP Description

000 command executed successfully

001-011 Reserved for future definition

100 command execution failed

101 command execution partially successful (Optional reasons)

110-111 Reserved for future definition

AdrType (3bit): NVE address type. Detailed definition as following.

AdrType Description

000 IPv4

001 IPv6

010-111 Reserved

4.2. VNTP Data Format

VNTP Data field varies according to different command.

For Register/Deregister Command, the field contains the NVE's address and a VN-ID set. In the VN-ID set, each entry for one VN-ID. See Figure 3.

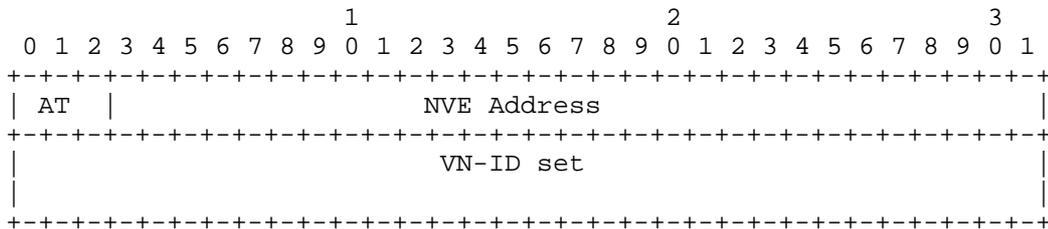


Figure 3 Data format for Register/Deregister command

Figure 3

The VN-ID set field would be Null, that means the NVE should be unreachable anymore.

For Request mapping info command, the field contains the Inner address set, each entry for one inner address. See figure 4.



Figure 4 Data format for Request command

Figure 4

The mapping/Address set field would be Null, that means NVA is requesting all the Inner address in this VN.

For Nullify Command, the field contains the NVE address. The command is used by NVA to notify all the NVEs in the same VN that the NVE is not reachable, all the mapping information relate to the NVE should be removed.

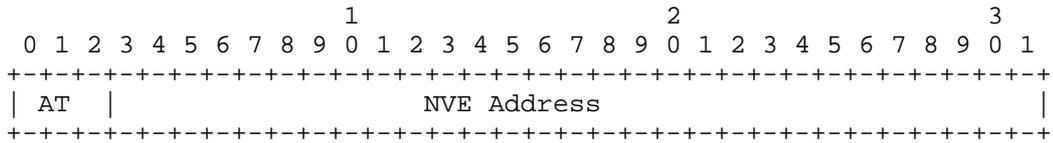


Figure 5 Data format for Nullify command

Figure 5

For Update Command, it also may include some entries, each entry has the detailed definition refers to Figure 3.

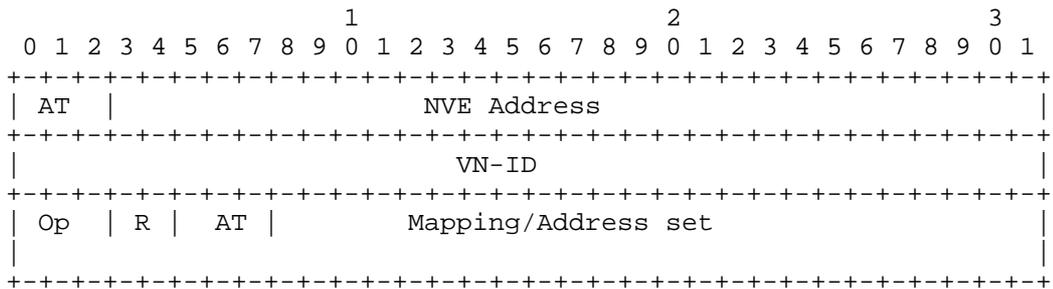


Figure 6 Data Format for Update command

Figure 6

The field means as following.

Op (3bit): for Operation Code.

Detailed definition for the Op field.

Op Description

000 Update add

001 Update delete

010 Update set to migration status

011 Update set to normal/non-migration status

100-111 Reserved for future use

AT (3bit): for Address Type.

Detailed definition for the AT field .

AT Description

000 IPv4

001 IPv6

010 MAC

011-111 Reserved

R (2bit): for Reserved.

NVE Address (length variable according AT): the outer address of mapping. If the update command is from NVE, the field is the local NVE address. If the update command is from NVA, the field is the remote NVE.

NV-ID: The Virtual Network ID that relate to the mapping information.

Mapping/Address (length variable according AT): each/inner Address needs updating. For the Update delete operation, if the field is Null, that means all the mapping info in this VN should be deleted.

5. The Operations of NVE

In the context of VNTP, the NVE works include:

1), If a VNI is created, the NVE will send Register command to NVA to register the VNI/NVE in the VN.

2), If a VNI is being deleted, the NVE will send update information to NVA to inform all the NVE related VN entry will be invalid. Or the NVA gets this information through the keep alive message, then nullify the all entries from this NVE_s VN.

3), If entries in the NVE have changed, for example, a new entry added or an existing entry deleted or become invalid, then the NVE will send update information to the NVA. Individual or batch update are supported.

4), And further, NVE also support tenant system migration.

5), The NVE accepts the updates from NVA and update the VRF table. The commands may be individual update or updates resulted from NVE failure.

6), Keep alive. Monitor the connection between NVE and NVA.

7), If the command not properly executed retransfer the command again for pre-setting times.

8), When NVA is unavailable or the NVA connection lost, optionally the NVE can connect other NVEs in the VN directly to keep the VN synchronized.

9), Security functions. TBD.

6. The operations of NVA

1), VNI creation

2), Form list of NVEs in the VN based on NVE Registration.

3), Accept updates from NVE and forward these updates to all other NVEs in the VN. Optionally, NVA store the update information for late use.

4), If NVE not register but update accepted, NVA may register it and forward the update to other NVEs.

5), if NVE registering after some updates then NVA will forward the stored updates to this NVE. Or NVA send request message to all other registered NVE for update if the previous updates not stored in NVA. And the NVA controls the updates only to this NVE other than all registered NVEs in the VN.

6), Keep alive. Monitor the connection between NVE and NVA.

7), if the command not properly executed NVA can retransfer the command again for pre-setting times.

8), When NVE in the VN is unavailable or the NVE connection lost, optionally the NVA can flood this NVE unreachable information to all other NVEs in the VN to keep the VN synchronized.

9), VNI delete. If there are not any VM or NVE in the VN, or the customer does not need the VN anymore then the NVA delete the VNI and release all the resources occupied by this VN.

10), Security functions. TBD.

7. Interaction with TS/Hypervisor-NVE protocol

Generally, VNTP can run independent of TS/Hypervisor-NVE protocol, but the interaction triggered by VRF changes because of the operation of TS/Hypervisor-NVE protocol. If the direct interaction is needed for further study.

8. Security Considerations

VNTP should support NVE and NVA mutual authentication and other security functions. The authentication has been covered by this draft, and the further security functions can be support through VNTP's command reservations.

9. IANA/IEEE Considerations

VNTP needs a specific IP protocol value, or TCP/UDP port allocation if the transport mechanism is chosen.

10. References

10.1. Normative references

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

- [RFC2234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, DOI 10.17487/RFC2234, November 1997, <<http://www.rfc-editor.org/info/rfc2234>>.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <<http://www.rfc-editor.org/info/rfc2328>>.
- [RFC5709] Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M., Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic Authentication", RFC 5709, DOI 10.17487/RFC5709, October 2009, <<http://www.rfc-editor.org/info/rfc5709>>.

10.2. Informative References

- [I-D.ietf-nvo3-arch]
Black, D., Hudson, J., Kreeger, L., Lasserre, M., and T. Narten, "An Architecture for Overlay Networks (NVO3)", draft-ietf-nvo3-arch-03 (work in progress), March 2015.
- [I-D.ietf-nvo3-hpvr2nve-cp-req]
Yizhou, L., Yong, L., Kreeger, L., Narten, T., and D. Black, "Split-NVE Control Plane Requirements", draft-ietf-nvo3-hpvr2nve-cp-req-03 (work in progress), August 2015.
- [I-D.ietf-nvo3-nve-nva-cp-req]
Kreeger, L., Dutt, D., Narten, T., and D. Black, "Network Virtualization NVE to NVA Control Protocol Requirements", draft-ietf-nvo3-nve-nva-cp-req-04 (work in progress), July 2015.
- [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, <<http://www.rfc-editor.org/info/rfc7364>>.
- [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, <<http://www.rfc-editor.org/info/rfc7365>>.

Authors' Addresses

Zhongyu Gu
Independent Reseacher

Email: guzhongyu2015@sina.com

Bhumip Khasnabish
ZTE (TX) Inc.
55 Madison Ave, Suite 302
Morristown, New Jersey 07960
USA

Phone: +001-781-752-8003
Email: vumip1@gmail.com, bhumip.khasnabish@ztetx.com
URI: <http://tinyurl.com/bhumip/>

Ting Ao
ZTE Corp
No.889 Bibo Rd
Shanghai
China

Phone: +86 21 68897642
Email: ao.ting@zte.com.cn

Fangwei Hu
ZTE Corp
No.889 Bibo Rd
Shanghai 201203
China

Phone: +86 21 68896273
Email: hu.fangwei@zte.com.cn