NVO3
Internet-Draft
Intended status: Standards Track
Expires: March 24, 2016

K. Singh
P. Jain
D. Garcia del Rio
Nuage Networks.
W. Henderickx
Alcatel-Lucent, Inc.
R. Shekhar
Juniper Networks
R. Rahman
Cisco Systems

September 21, 2015

VxLAN Router Alert Option
draft-singh-nvo3-vxlan-router-alert-02

Abstract

This proposal describes a new option to achieve a mechanism which alerts VxLAN terminating VTEP to more closely examine the contents of the packet encapsulated under VxLAN header. This option is useful for case(s) where a given frame encapsulated within a given VXLAN segment responsible for carrying data between two different End Systems contains some control information (e.g OAM information, any control plane protocol packet etc.) that may require special handling/processing by terminating VTEP.

Status of this Memo

Copyright Notice

Table of Contents

1. 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 RFC2119 [RFC2119].

When used in lower case, these words convey their typical use in common language, and are not to be interpreted as described in RFC2119 [RFC2119].

2. Introduction

VXLAN [RFC7348]is a tunneling mechanism to overlay Layer 2 networks on top of Layer 3 networks. In most cases the end point of the tunnel (VTEP) is intended to be at the edge of the network, typically connecting an access switch to an IP transport network. The access switch could be a physical or a virtual switch located within the hypervisor on the server which is connected to End System which is a VM.

VXLAN segment encapsulates End System data at Originating VTEP and carries it over L3 network to the Terminating VTEP, where VXLAN header is interpreted, removed and data is passed on to the End System.

There could be some scenarios, where the network element at originating VTEP needs to encapsulate some control information in a given VXLAN segment, and this control information needs to be analysed and processed at the terminating VTEP for that VXLAN segment. There could be various examples of such control information e.g OAM, and protocol control packets encapsulated in VxLAN segment.

This document defines a mechanism whereby Originating VTEP can add additional information to the VXLAN header, based upon which the Terminating VTEP can decide to analyse the payload under VXLAN packet and handle it to slow-path, rather then forwarding it to the destination End System.

3. Terminology

Terminology used in this document:

VXLAN: Virtual eXtensible Local Area Network.

VTEP: VXLAN Tunnel End Point.

VM: Virtual Machine.

End System: Could be VM etc. - System whose data is expected to go over VXLAN Segment.

OAM: Operations, Administration, and Maintenance

Other terminologies are as defined in [RFC7348].

4. Approach

If the Originating VTEP decides to generate control information, which needs to go over a given VXLAN segment and if the Terminating VTEP needs to analyse and process it, then following procedures have to be followed at Originating and Terminating VTEP(s):-

4.1. Originating VTEP Procedure

When creating the VXLAN header for a given VXLAN segment, the Originating VTEP MUST set Router Alert Bit in the Flag bits in VXLAN header. The VNI for this frame MUST be the same as for the given VXLAN segment which carries the data traffic of the End System.

```
    VXLAN Header:
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |R|R|R|R|I|R|R|RA|           Reserved                          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                VXLAN Network Identifier (VNI) |   Reserved    |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

    RA: Router Alert Bit (Proposed)
```

4.2. Terminating VTEP Procedure

On receiving VXLAN frame, the Terminating VTEP would do the usual VXLAN processing as defined in [RFC7348], but if the RA Bit in Flags is Set it MUST send the rest of the inner frame for further processing to the above application. The details of the applications and how it would process the inner frame is outside the scope of this document. This frame MUST not be sent to the target End System.

5. Management Considerations

None

6. Security Considerations

For wide range, security requirements for VxLAN Packets with Route Alert (RA) bit set is no different from how IP Router Alert Option is handled in Network End Points.

The most common potential attack could be Denial-of-Service attacks by sending VxLAN Packets with Router Alert Bit Set at aggressive rate, causing potential high resource utilization. For such scenarios its recommended that implementations regulate sending of such packets to control plane via rate limiting.

7. Acknowledgements

The authors want to thank Krishna Ram Kuttuva Jeyaram, and Suresh Boddapati of Nuage Networks for significant contribution and feedback.

8. IANA Considerations

Router Alert Bit (RA): IANA is request to asigh 1 Bit in Flags field of VXLAN Header to communicate VXLAN Router Alert information.

9. References

9.1. Normative References

[I-D.draft-lasserre-nvo3-framework]
Lasserre, M., Balus, F., Morin, T., Bitar, N., and Y. Rekhter, "Framework for DC Network Virtualization", September 2011.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <http://www.rfc-editor.org/info/rfc7348>.

9.2. Informative 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>.

Authors' Addresses

Kanwar Singh
Nuage Networks.
755 Ravendale Drive
Mountain View, CA 94043
USA

Email: kanwar@nuagenetworks.net

Pradeep Jain
Nuage Networks.
755 Ravendale Drive
Mountain View, CA 94043
USA

Email: pradeep@nuagenetworks.net

Diego
Nuage Networks.
755 Ravendale Drive
Mountain View, CA 94043
USA

Email: diego@nuagenetworks.net

Wim Henderickx
Alcatel-Lucent, Inc.
Copernicuslaan 50, 2018
ANTWERP 2018
BELGIUM

Email: Wim.Henderickx@alcatel-lucent.com

Ravi Shekhar
Juniper Networks
1194 North Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: rshekhar@juniper.net

Reshad Rahman
Cisco Systems
2000 Innovation Drive
Kanata, ON K2K 3E8
USA

Email: rrahman@cisco.com