

NVO3 Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 11, 2017

G. Mirsky
ZTE Corp.
N. Kumar
D. Kumar
Cisco Systems, Inc.
M. Chen
Y. Li
Huawei Technologies
D. Dolson
Sandvine
March 10, 2017

Echo Request and Echo Reply for Overlay Networks
draft-ooamdt-rtgwg-demand-cc-cv-03

Abstract

This document defines Overlay Echo Request and Echo Reply that enable on-demand Continuity Check, Connectivity Verification among other operations in overlay networks.

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 September 11, 2017.

Copyright Notice

Copyright (c) 2017 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
 - 1.1. Conventions used in this document 2
 - 1.1.1. Terminology 2
 - 1.1.2. Requirements Language 3
- 2. On-demand Continuity Check and Connectivity Verification . . 3
 - 2.1. Requirements Towards On-demand CC/CV OAM 3
 - 2.2. Proposed Solution 4
 - 2.3. Overlay Echo Request Transmission 5
 - 2.4. Overlay Echo Request Reception 6
 - 2.5. Overlay Echo Reply Transmission 6
 - 2.6. Overlay Echo Reply Reception 6
- 3. IANA Considerations 6
 - 3.1. Overlay Echo Request/Echo Reply Type 6
 - 3.2. Overlay Ping Parameters 6
 - 3.3. Overlay Echo Request/Echo Reply Message Types 6
 - 3.4. Overlay Echo Reply Modes 7
- 4. Security Considerations 7
- 5. Contributors 8
- 6. Acknowledgment 9
- 7. References 9
 - 7.1. Normative References 9
 - 7.2. Informative References 10
- Authors' Addresses 10

1. Introduction

Operations, Administration, and Maintenance (OAM) toolset provides methods for fault management and performance monitoring in each layer of the network, in order to improve their ability to support services with guaranteed and strict Service Level Agreements (SLAs) while reducing operational costs.

1.1. Conventions used in this document

1.1.1. Terminology

Term "Overlay OAM" used in this document interchangeably with longer version "set of OAM protocols, methods and tools for Overlay networks". And "Overlay ping" is used interchangeably with longer version Overlay Echo Request/Reply.

CC Continuity Check

CV Connectivity Verification

ECMP Equal Cost Multipath

FM Fault Management

Geneve Generic Network Virtualization Encapsulation

GUE Generic UDP Encapsulation

MPLS Multiprotocol Label Switching

NVO3 Network Virtualization Overlays

OAM Operations, Administration, and Maintenance

SFC Service Function Chaining

SFP Service Function Path

VXLAN Virtual eXtensible Local Area Network

VXLAN-GPE Generic Protocol Extension for VXLAN

1.1.2. 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 [RFC2119].

2. On-demand Continuity Check and Connectivity Verification

2.1. Requirements Towards On-demand CC/CV OAM

Availability, not as performance metric, is understood as ability to reach the node, i.e. the fact that path between ingress and egress does exist. Such OAM mechanism also referred as Continuity Check (CC). Connectivity Verification (CV) extends Continuity Check functionality in order to provide confirmation that the desired source is connected to the desired sink.

Echo Request/Reply OAM mechanism enables detection of the loss of continuity defect, its localization and collection information in order to discover root cause. These are requirements considered:

REQ#1: MUST support fault localization of Loss of Continuity check at Overlay layer.

REQ#2: MAY support fault localization of Loss of Continuity check at transport layer.

REQ#3: MUST support tracing path in overlay network through the overlay nodes.

REQ#4: MAY support tracing path in underlay network connecting overlay border nodes.

REQ#5: MAY support verification of the mapping between its data plane state and client layer services.

REQ#6: MUST have the ability to discover and exercise equal cost multipath (ECMP) paths in its underlay network.

REQ#7: MUST be able to trigger on-demand FM with responses being directed towards initiator of such proxy request.

2.2. Proposed Solution

The format of the Echo Request/Echo Reply control packet is to support ping and traceroute functionality in overlay networks. Figure 1 resembles the format of MPLS LSP Ping [RFC4379] with some exceptions.

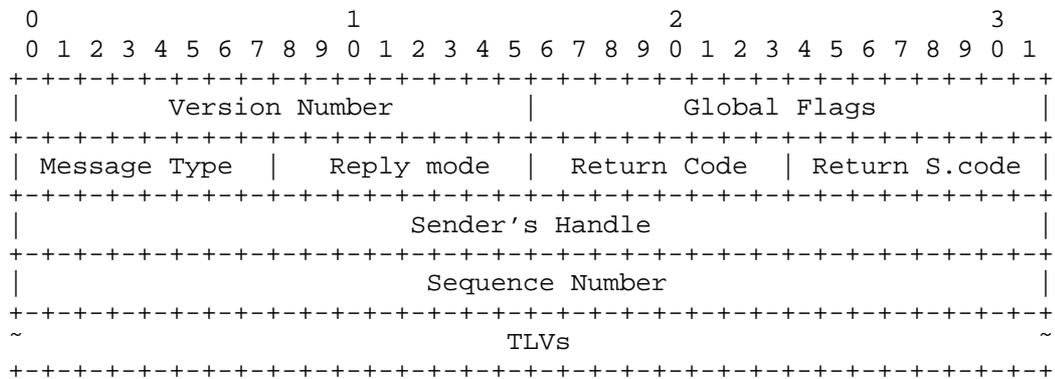


Figure 1: Overlay OAM Ping format

The interpretation of the fields is as following:

The Version reflects the current version. The version number is to be incremented whenever a change is made that affects the

ability of an implementation to correctly parse or process control packet.

The Global Flags is a bit vector field

The Message Type field reflects the type of the packet. Value TBA2 identifies Echo Request and TBA3 - Echo Reply

The Reply Mode defines the type of the return path requested by the sender of the Echo Request.

Return Codes and Subcodes can be used to inform the sender about result of processing its request.

The Sender's Handle is filled in by the sender, and returned unchanged by the receiver in the Echo Reply.

The Sequence Number is assigned by the sender and can be (for example) used to detect missed replies.

TLVs (Type-Length-Value tuples) have the two octets long Type field, two octets long Length field that is length of the Value field in octets.

2.3. Overlay Echo Request Transmission

Overlay Echo Request control packet MUST use the appropriate encapsulation of the monitored overlay network. Overlay network encapsulation MUST identify Echo Request as OAM packet. Overlay encapsulation uses different methods to identify OAM payload [I-D.ietf-nvo3-vxlan-gpe], [I-D.ietf-nvo3-gue], [I-D.ietf-nvo3-geneve], [I-D.ietf-sfc-nsh], [I-D.ietf-bier-mpls-encapsulation]. Overlay network's header MUST be immediately followed by the Overlay OAM Header [I-D.ooamdt-rtgwg-ooam-header]. Message Type field in the Overlay OAM Header MUST be set to Overlay Echo Request value (TBA2).

Value of the Reply Mode field MAY be set to:

- o Do Not Reply (TBA4) if one-way monitoring is desired. If Echo Request is used to measure synthetic packet loss, the receiver MAY report loss measurement results to a remote node.
- o Reply via an IPv4/IPv6 UDP Packet (TBA5) value likely will be the most used.
- o Reply via Application Level Control Channel (TBA6) value if the overlay network MAY have bi-directional paths.

- o Reply via Specified Path (TBA7) value in order to enforce use of the particular return path specified in the included TLV to verify bi-directional continuity and also increase robustness of the monitoring by selecting more stable path.

2.4. Overlay Echo Request Reception

2.5. Overlay Echo Reply Transmission

The Reply Mode field directs whether and how the Echo Reply message should be sent. The sender of the Echo Request MAY use TLVs to request that corresponding Echo Reply be sent using the specified path. Value TBA3 is referred as "Do not reply" mode and suppresses transmission of Echo Reply packet. Default value (TBA5) for the Reply mode field requests the responder to send the Echo Reply packet out-of-band as IPv4 or IPv6 UDP packet. [Selection of destination and source IP addresses and UDP port numbers to be provided in the next update.]

2.6. Overlay Echo Reply Reception

3. IANA Considerations

3.1. Overlay Echo Request/Echo Reply Type

IANA is requested to assign new type from the Overlay OAM Protocol Types registry as follows:

Value	Description	Reference
TBA1	Overlay Echo Request/Echo Reply	This document

Table 1: Overlay Echo Request/Echo Reply Type

3.2. Overlay Ping Parameters

IANA is requested to create new Overlay Echo Request/Echo Reply Parameters registry.

3.3. Overlay Echo Request/Echo Reply Message Types

IANA is requested to create in the Overlay Echo Request/Echo Reply Parameters registry the new sub-registry Message Types. All code points in the range 1 through 191 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC5226] and assign values as follows:

Value	Description	Reference
0	Reserved	
TBA2	Overlay Echo Request	This document
TBA3	Overlay Echo Reply	This document
TBA3+1-191	Unassigned	IETF Review
192-251	Unassigned	First Come First Served
252-254	Unassigned	Private Use
255	Reserved	

Table 2: Overlay Echo Request/Echo Reply Message Types

3.4. Overlay Echo Reply Modes

IANA is requested to create in the Overlay Echo Request/Echo Reply Parameters registry the new sub-registry Reply Modes All code points in the range 1 through 191 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC5226] and assign values as follows:

Value	Description	Reference
0	Reserved	
TBA4	Do Not Reply	This document
TBA5	Reply via an IPv4/IPv6 UDP Packet	This document
TBA6	Reply via Application Level Control Channel	This document
TBA7	Reply via Specified Path	This document
TBA7+1-191	Unassigned	IETF Review
192-251	Unassigned	First Come First Served
252-254	Unassigned	Private Use
255	Reserved	

Table 3: Overlay Echo Reply Modes

4. Security Considerations

Overlay Echo Request/Reply operates within the domain of the overlay network and thus inherits any security considerations that apply to the use of that overlay technology and, consequently, underlay data plane. Also, the security needs for Overlay Echo Request/Reply are

similar to those of ICMP ping [RFC0792], [RFC4443] and MPLS LSP ping [I-D.ietf-mpls-rfc4379bis].

There are at least three approaches of attacking a node in the overlay network using the mechanisms defined in the document. One is a Denial-of-Service attack, by sending Overlay ping to overload a node in the overlay network. The second may use spoofing, hijacking, replying, or otherwise tampering with Overlay Echo Requests and/or Replies to misrepresent, alter operator's view of the state of the overlay network. The third is an unauthorized source using an Overlay Echo Request/Reply to obtain information about the overlay and/or underlay network.

To mitigate potential Denial-of-Service attacks, it is RECOMMENDED that implementations throttle the Overlay ping traffic going to the control plane.

Reply and spoofing attacks involving faking or replying Overlay Echo Reply messages would have to match the Sender's Handle and Sequence Number of an outstanding Overlay Echo Request message which is highly unlikely. Thus the non-matching reply would be discarded. But since "even a broken clock is right twice a day" implementations MAY use Timestamp control block [I-D.ooamdt-rtgwg-ooam-header] to validate the TimeStamp Sent by requiring an exact match on this field.

To protect against unauthorized sources trying to obtain information about the overlay and/or underlay an implementation MAY check that the source of the Echo Request is indeed part of the overlay domain.

5. Contributors

Work on this document started by Overlay OAM Design Team with contributions from:

Carlos Pignataro

Cisco Systems, Inc.

cpignata@cisco.com

Santosh Pallagatti

santosh.pallagatti@gmail.com

Erik Nordmark

Arista Networks

nordmark@acm.org

Ignas Bagdonas

ibagdona@gmail.com

David Mozes

Mellanox Technologies Ltd.

davidm@mellanox.com

6. Acknowledgment

TBD

7. References

7.1. Normative References

[I-D.ietf-bier-mpls-encapsulation]

Wijnands, I., Rosen, E., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication in MPLS and non-MPLS Networks", draft-ietf-bier-mpls-encapsulation-06 (work in progress), December 2016.

[I-D.ietf-nvo3-geneve]

Gross, J., Ganga, I., and T. Sridhar, "Geneve: Generic Network Virtualization Encapsulation", draft-ietf-nvo3-geneve-03 (work in progress), September 2016.

[I-D.ietf-nvo3-gue]

Herbert, T., Yong, L., and O. Zia, "Generic UDP Encapsulation", draft-ietf-nvo3-gue-05 (work in progress), October 2016.

[I-D.ietf-nvo3-vxlan-gpe]

Maino, F., Kreeger, L., and U. Elzur, "Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-03 (work in progress), October 2016.

[I-D.ietf-sfc-nsh]

Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-12 (work in progress), February 2017.

- [I-D.ooamdt-rtgwg-ooam-header]
Mirsky, G., Kumar, N., Kumar, D., Chen, M., Yizhou, L.,
Mozes, D., and D. Dolson, "OAM Header for use in Overlay
Networks", draft-ooamdt-rtgwg-ooam-header-02 (work in
progress), February 2017.
- [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>>.

7.2. Informative References

- [I-D.ietf-mpls-rfc4379bis]
Kompella, K., Swallow, G., Pignataro, C., Kumar, N.,
Aldrin, S., and M. Chen, "Detecting Multi-Protocol Label
Switched (MPLS) Data Plane Failures", draft-ietf-mpls-
rfc4379bis-09 (work in progress), October 2016.
- [RFC0792] Postel, J., "Internet Control Message Protocol", STD 5,
RFC 792, DOI 10.17487/RFC0792, September 1981,
<<http://www.rfc-editor.org/info/rfc792>>.
- [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol
Label Switched (MPLS) Data Plane Failures", RFC 4379,
DOI 10.17487/RFC4379, February 2006,
<<http://www.rfc-editor.org/info/rfc4379>>.
- [RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet
Control Message Protocol (ICMPv6) for the Internet
Protocol Version 6 (IPv6) Specification", RFC 4443,
DOI 10.17487/RFC4443, March 2006,
<<http://www.rfc-editor.org/info/rfc4443>>.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an
IANA Considerations Section in RFCs", BCP 26, RFC 5226,
DOI 10.17487/RFC5226, May 2008,
<<http://www.rfc-editor.org/info/rfc5226>>.

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Nagendra Kumar
Cisco Systems, Inc.

Email: naikumar@cisco.com

Deepak Kumar
Cisco Systems, Inc.

Email: dekumar@cisco.com

Mach Chen
Huawei Technologies

Email: mach.chen@huawei.com

Yizhou Li
Huawei Technologies

Email: liyizhou@huawei.com

David Dolson
Sandvine

Email: ddolson@sandvine.com