Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: August 22, 2013

X. Chen
Huawei Technologies
T. Tsou
Huawei Technologies (USA)
E. Roch
Huawei Technologies

February 18, 2013

NVO3 Requirements Versus Available Protocol Capabilities
draft-chen-nvo3-gap-analysis-00

Abstract

This document matches candidate protocols against the NVO3 requirements. Based on the results, gaps are identified and further protocol work is recommended.

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 August 22, 2013.

Table of Contents

1. Introduction

The charter of the NVO3 Working Group requires it to identify any gaps between the requirements it has identified and the available protocol solutions as a prerequisite to rechartering or concluding the Working Group if no gaps exist. The present document is intended to provide the required analysis. It provides a tabulation of the candidate protocols' ability to satisfy each requirement identified by the Working Group. Areas where further work is required to ensure that the requirements are met are identified.

Since the Working Group has yet to adopt documents describing requirements for the management and control planes, they are absent from the present version of this document. The data plane requirements are taken from [I_D.dataplane_requirements]. The initial candidate protocols are NVGRE [I_D.NVGRE], VxLAN [I_D.VxLAN], L2VPN [reference?], and L3VPN [reference?].

1.1. Requirements Language

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].

1.2. Abbreviations

This document uses the following abbreviations:

NVO3: Network virtualization overlays

L2VPN Layer 2 virtual private network

L3VPN Layer 3 virtual private network

NVE: Network virtualization edge

VAP: Virtual access point

VNI: Virtual network instance

LAG: Link aggregation group

ECMP: Equal cost multi-path

DSCP: Differentiated services code point

ECN: Explicit congestion notification [RFC3168]

2. Management Requirements

To come.

3. Control Plane Requirements

To come.

4. Data Plane Requirements

In this section, the numbering of requirement headings is taken from the corresponding section numbers in [I_D.dataplane_requirements].

3.1. Virtual Access Points (VAPs)

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
MUST support VAP identification				
-	-	-	-	-
1) Local interface	YES			
-	-	-	-	-
2) Local interface + fields in frame header	YES			

Table 1: VAP Identification Requirements

3.2. Virtual Network Instance (VNI)

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
VAP are associated with a specific VNI at service instantiation time.	YES			

Table 2: VAP-VNI Association

3.2.1. L2 VNI

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
L2 VNI MUST provide an emulated Ethernet multipoint service as if Tenant Systems are interconnected by a bridge (but instead by using a set of NVO3 tunnels).	NO			
-	-	-	-	-
Loop avoidance capability MUST be provided.				
-	-	-	-	-
In the absence of a management or control plane, data plane learning MUST be used to populate forwarding tables.				
-	-	-	-	-
When flooding is required, either to deliver unknown unicast, or broadcast or multicast traffic, the NVE MUST either support ingress replication or multicast.				
-	-	-	-	-
In this latter case, the NVE MUST be able to build at least a default flooding tree per VNI.				

Table 3: L2 VNI Service

3.2.2. L3 VNI

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
L3 VNIs MUST provide virtualized IP routing and forwarding.	YES			
-	-	-	-	-

L3 VNIs MUST support per-tenant forwarding instance with IP addressing isolation and L3 tunneling for interconnecting instances of the same VNI on NVEs.	YES			

Table 4: L3 VNI Service

3.3.1. NVO3 overlay header

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
An NVO3 overlay header MUST be included after the underlay tunnel header when forwarding tenant traffic.	YES			

Table 5: Overlay Header

3.3.1.1. Virtual Network Context Identification

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
The overlay encapsulation header MUST contain a field which allows the encapsulated frame to be delivered to the appropriate virtual network endpoint by the egress NVE.	YES			

Table 6: Virtual Network Context Identification

3.3.1.2. Service QoS identifier

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
Traffic flows originating from different applications could rely on differentiated forwarding treatment to meet end-to-end availability and performance objectives.	NO			

Table 7: QoS Service Identification

3.3.2.1. LAG and ECMP

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
For performance reasons, multipath over LAG and ECMP paths SHOULD be supported.	YES			

Table 8: Multipath Support

3.3.2.2. DiffServ and ECN marking

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
[RFC2983] defines two modes for mapping the DSCP markings from inner to outer headers and vice versa. Both models SHOULD be supported.	NO			
-	-	-	-	-
ECN marking MUST be performed according to [RFC6040] which describes the correct ECN behavior for IP tunnels.	NO			

Table 9: DSCP and ECN Marking

3.3.2.3. Handling of broadcast, unknown unicast, and multicast traffic

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
NVO3 data plane support for either ingress replication or point-to- multipoint tunnels is required to send traffic destined to multiple locations on a per-VNI basis (e.g. L2/L3 multicast traffic, L2 broadcast and unknown unicast traffic).	YES			

Table 10: Handling of Broadcast, Unknown Unicast, and Multicast Traffic

3.4. External NVO3 connectivity

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
NVO3 services MUST interoperate with current VPN and Internet services. This may happen inside one DC during a migration phase or as NVO3 services are delivered to the outside world via Internet or VPN gateways.	YES			

Table 11: Interoperation

3.5. Path MTU

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
Classical ICMP-based MTU Path Discovery ([RFC1191], [RFC1981]) or Extended MTU Path Discovery techniques such as defined in [RFC4821].	NO			
-	-	-	-	-

Segmentation and reassembly support from the overlay layer operations without relying on the Tenant Systems to know about the end-to-end MTU.	YES			

Table 12: Path MTU

3.7. NVE Multi-Homing Requirements

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
Multi-homing techniques SHOULD be used to increase the reliability of an NVO3 network.	NO			

Table 13: Multihoming

3.8. OAM

Requirement	NVGRE	VxLAN	L2VPN	L3VPN
NVE MAY be able to originate/terminate OAM messages for connectivity verification, performance monitoring, statistic gathering and fault isolation. Depending on configuration, NVEs SHOULD be able to process or transparently tunnel OAM messages, as well as supporting alarm propagation capabilities.	NO			

Table 14: OAM Messaging

5. Summary and Conclusions

To come.

6. Acknowledgements

Peter Ashwood-Smith and Rangaraju Iyengar are acknowledged for their technical contributions to this document. Tom Taylor served as XML2RFC guru to produce it.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

All drafts are required to have a security considerations section.

9. References

9.1. Normative References

[I_D.NVGRE] Sridharan, M., Greenberg, A., Venkataramiah, N., Wang, Y., Duda, K., Ganga, I., Lin, G., Pearson, M., Thaler, P., and C. Tumuluri, "NVGRE: Network Virtualization using Generic Routing Encapsulation (Work in progress)", July 2012.

[I_D.VxLAN] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks (Work in progress)", August 2012.

[I_D.dataplane_requirements] Bitar, N., Lasserre, M., Balus, F., Morin, T., Jin, L., and B. Khasnabish, "NVO3 Data Plane Requirements (Work in progress)", December 2012.

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2983] Black, D., "Differentiated Services and Tunnels", RFC 2983, October 2000.

[RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", RFC 4821, March 2007.

[RFC6040] Briscoe, B., "Tunnelling of Explicit Congestion Notification", RFC 6040, November 2010.

9.2. Informative References

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, September 2001.

Authors' Addresses

Xiaoming Chen
Huawei Technologies

Phone:
Email: ming.chen@huawei.com

Tina Tsou
Huawei Technologies (USA)
2330 Central Expressway
Santa Clara, CA 95050
USA

Phone: +1 408 330 4424
Email: Tina.Tsou.Zouting@huawei.com
URI: http://tinatsou.weebly.com/contact.html

Evelyne Roch
Huawei Technologies
303 Terry Fox Drive, Suite 400
Kanata, Ontario K2K 3J1
Canada

Phone: +1 613 595 1900 x1612
Email: evelyne.roch@huawei.com