        Use Cases for Data Center Network Virtualization Overlay Networks

                    draft-ietf-nvo3-use-case-17

Abstract

   This document describes data center network virtualization overlay
   (NVO3) network use cases that can be deployed in various data
   centers and serve different data center applications.

Status of this Memo

Table of Contents

[1](#). Introduction

   Server virtualization has changed the Information Technology (IT)
   industry in terms of the efficiency, cost, and speed of providing
   new applications and/or services such as cloud applications. However
   traditional data center (DC) networks have limits in supporting
   cloud applications and multi tenant networks [[RFC7364](#)]. The goals of
   data center network virtualization overlay (NVO3) networks are to
   decouple the communication among tenant systems from DC physical
   infrastructure networks and to allow one physical network
   infrastructure to:

   o Carry many NVO3 virtual networks and isolate the traffic of
      different NVO3 virtual networks on a physical network.

   o Provide independent address space in individual NVO3 virtual
      network such as MAC and IP.

   o Support flexible Virtual Machines (VM) and/or workload placement
      including the ability to move them from one server to another
      without requiring VM address changes and physical infrastructure
      network configuration changes, and the ability to perform a "hot
      move" with no disruption to the live application running on those
      VMs.

   These characteristics of NVO3 virtual networks help address the
   issues that cloud applications face in data centers [[RFC7364](#)].

   Hosts in one NVO3 virtual network may communicate with hosts in
   another NVO3 virtual network that is carried by the same physical
   network, or different physical network, via a gateway. The use case
   examples for the latter are: 1) DCs that migrate toward an NVO3
   solution will be done in steps, where a portion of tenant systems in
   a VN are on virtualized servers while others exist on a LAN. 2) many
   DC applications serve to Internet users who are on different
   physical networks; 3) some applications are CPU bound, such as Big
   Data analytics, and may not run on virtualized resources. The inter-
   VN policies are usually enforced by the gateway.

   This document describes general NVO3 virtual network use cases that
   apply to various data centers. The use cases described here
   represent DC provider's interests and vision for their cloud
   services. The document groups the use cases into three categories
   from simple to sophiscated in terms of implementation. However the
   implementation details of these use cases are outside the scope of
   this document. These three categories are highlighted below:

o Basic NVO3 virtual networks (Section 2). All Tenant Systems (TS)
   in the network are located within the same DC. The individual
   networks can be either Layer 2 (L2) or Layer 3 (L3). The number
   of NVO3 virtual networks in a DC is much larger than the number
   that traditional VLAN based virtual networks [IEEE 802.1Q] can
   support.

o A virtual network that spans across multiple Data Centers and/or
   to customer premises where NVO3 virtual networks are constructed
   and interconnect other virtual or physical networks outside the
   data center. An enterprise customer may use a traditional
   carrier-grade VPN or an IPsec tunnel over the Internet to
   communicate with its systems in the DC. This is described in
   Section 3.

o DC applications or services require an advanced network that
   contains several NVO3 virtual networks that are interconnected by
   gateways. Three scenarios are described in Section 4. (1)
   supporting multiple technologies; (2) constructing several
   virtual networks as a tenant network; (3) applying NVO3 to a
   virtual Data Center (vDC).

The document uses the architecture reference model defined in
[RFC7365] to describe the use cases.

1.1. Terminology

This document uses the terminology defined in [RFC7365] and
[RFC4364]. Some additional terms used in the document are listed
here.

ASBR: Autonomous System Border Routers (ASBR)

DMZ: Demilitarized Zone. A computer or small sub-network that sits
between a more trusted internal network, such as a corporate private
LAN, and an un-trusted or less trusted external network, such as the
public Internet.

DNS: Domain Name Service [RFC1035]

DC Operator: An entity that is responsible for constructing and
managing all resources in data centers, including, but not limited
to, compute, storage, networking, etc.

DC Provider: An entity that uses its DC infrastructure to offer
services to its customers.

NAT: Network Address Translation [RFC3022]

vGW: virtual Gateway; a gateway component used for an NVO3 virtual
network to interconnect with another virtual/physical network.

NVO3 virtual network: a virtual network that is implemented based
NVO3 architecture [NVO3-ARCH].

PE: Provider Edge

SP: Service Provider

TS: A TS can be a physical server/device or a virtual machine (VM)
on a server, i.e., end-device [RFC7365].

VRF-LITE: Virtual Routing and Forwarding - LITE [VRF-LITE]

VN: NVO3 virtual network.

WAN VPN: Wide Area Network Virtual Private Network [RFC4364]
[RFC7432]

1.2. NVO3 Background

An NVO3 virtual network is a virtual network in a DC that is
implemented based on the NV03 architecture [RFC8014]. This
architecture is often referred to as an overlay architecture. The
traffic carried by an NVO3 virtual network is encapsulated at a
Network Virtual Edge (NVE) [RFC8014] and carried by a tunnel to
another NVE where the traffic is decapsulated and sent to a
destination Tenant System (TS). The NVO3 architecture decouples NVO3
virtual networks from the DC physical network configuration. The
architecture uses common tunnels to carry NVO3 traffic that belongs
to multiple NVO3 virtual networks.

An NVO3 virtual network may be an L2 or L3 domain. The network
provides switching (L2) or routing (L3) capability to support host
(i.e., tenant systems) communications. An NVO3 virtual network may
be required to carry unicast traffic and/or multicast,
broadcast/unknown-unicast (for L2 only) traffic from/to tenant
systems. There are several ways to transport NVO3 virtual network
BUM (Broadcast, Unknown-unicast, Multicast) traffic [NVO3MCAST].

An NVO3 virtual network provides communications among Tenant Systems
(TS) in a DC. A TS can be a physical server/device or a virtual
machine (VM) on a server end-device [RFC7365].

2. **DC with Large Number of Virtual Networks**

   A DC provider often uses NVO3 virtual networks for internal
   applications where each application runs on many VMs or physical
   servers and the provider requires applications to be segregated from
   each other. A DC may run a larger number of NVO3 virtual networks to
   support many applications concurrently, where traditional IEEE802.1Q
   based VLAN solution is limited to 4094 VLANs.

   Applications running on VMs may require different quantity of
   computing resource, which may result in computing resource shortage
   on some servers and other servers being nearly idle. Shortage of
   computing resource may impact application performance. DC operators
   desire VM or workload movement for resource usage optimization. VM
   dynamic placement and mobility results in frequent changes of the
   binding between a TS and an NVE. The TS reachability update
   mechanisms should take significantly less time than the typical re-
   transmission Time-out window of a reliable transport protocol such
   as TCP and SCTP, so that end points' transport connections won't be
   impacted by a TS becoming bound to a different NVE. The capability
   of supporting many TSs in a virtual network and many virtual
   networks in a DC is critical for an NVO3 solution.

   When NVO3 virtual networks segregate VMs belonging to different
   applications, DC operators can independently assign MAC and/or IP
   address space to each virtual network. This addressing is more
   flexible than requiring all hosts in all NVO3 virtual networks to
   share one address space. In contrast, typical use of IEEE 802.1Q
   VLANs requires a single common MAC address space.

3. **DC NVO3 virtual network and External Network Interconnection**

   Many customers (enterprises or individuals) who utilize a DC
   provider's compute and storage resources to run their applications
   need to access their systems hosted in a DC through Internet or
   Service Providers' Wide Area Networks (WAN). A DC provider can
   construct a NVO3 virtual network that provides connectivity to all
   the resources designated for a customer and allows the customer to
   access the resources via a virtual gateway (vGW). WAN connectivity
   to the virtual gateway can be provided by VPN technologies such as
   IPsec VPNs [RFC4301] and BGP/MPLS IP VPNs [RFC 4364].

   If a virtual network spans multiple DC sites, one design using NVO3
   is to allow the network to seamlessly span the sites without DC
   gateway routers' termination. In this case, the tunnel between a

   pair of NVEs can be carried within other intermediate tunnels over
   the Internet or other WANs, or an intra-DC tunnel and inter DC
   tunnel(s) can be stitched together to form an end-to-end tunnel
   between the pair of NVEs that are in different DC sites. Both cases
   will form one NVO3 virtual network across multiple DC sites.

   Two use cases are described in the following sections.

3.1. DC NVO3 virtual network Access via the Internet

   A customer can connect to an NVO3 virtual network via the Internet
   in a secure way. Figure 1 illustrates an example of this case. The
   NVO3 virtual network has an instance at NVE1 and NVE2 and the two
   NVEs are connected via an IP tunnel in the Data Center. A set of
   tenant systems are attached to NVE1 on a server. NVE2 resides on a
   DC Gateway device. NVE2 terminates the tunnel and uses the VNID on
   the packet to pass the packet to the corresponding vGW entity on the
   DC GW (the vGW is the default gateway for the virtual network). A
   customer can access their systems, i.e., TS1 or TSn, in the DC via
   the Internet by using an IPsec tunnel [RFC4301]. The IPsec tunnel is
   configured between the vGW and the customer gateway at the customer
   site. Either a static route or Interior Border Gateway Protocol
   (iBGP) may be used for prefix advertisement. The vGW provides IPsec
   functionality such as authentication scheme and encryption; iBGP
   protocol traffic is carried within the IPsec tunnel. Some vGW
   features are listed below:

   o The vGW maintains the TS/NVE mappings and advertises the TS
      prefix to the customer via static route or iBGP.

   o Some vGW functions such as firewall and load balancer can be
      performed by locally attached network appliance devices.

   o If the NVO3 virtual network uses different address space than
      external users, then the vGW needs to provide the NAT function.

   o More than one IPsec tunnel can be configured for redundancy.

   o The vGW can be implemented on a server or VM. In this case, IP
      tunnels or IPsec tunnels can be used over the DC infrastructure.

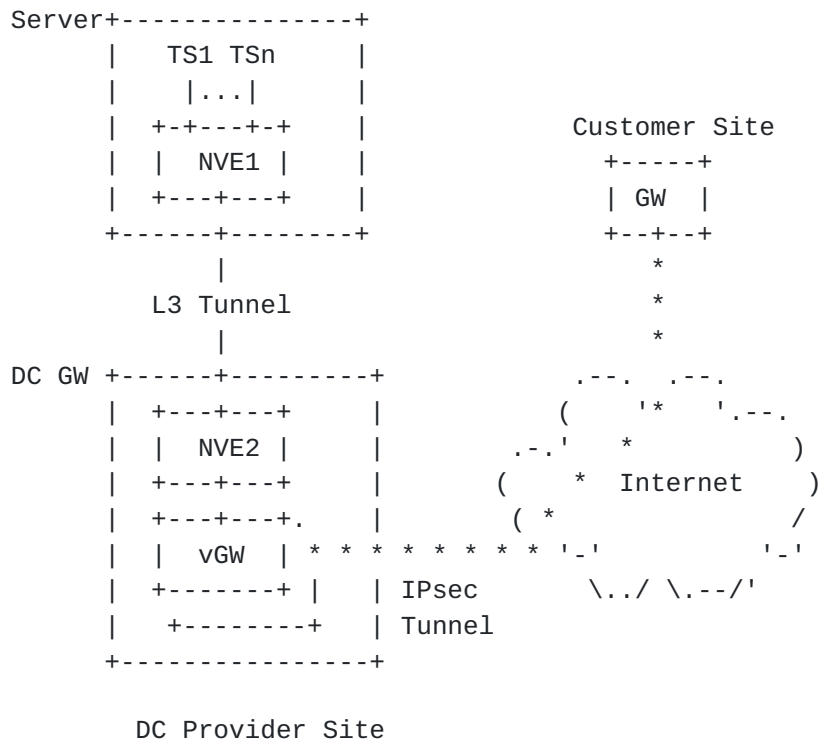   o DC operators need to construct a vGW for each customer.

```
   Server+----------------+
         |   TS1 TSn      |
         |   |...|        |
         |  +-+---+-+     |                Customer Site
         |  | NVE1 |      |                   +-----+
         |  +---+---+     |                   | GW  |
         +------+--------+                    +--+--+
                |                                *
          L3 Tunnel                              *
                |                                *
   DC GW +------+---------+             .--.   .--.
         | +---+---+      |            (     '*    '.--.
         | | NVE2 |       |         .-.'    *          )
         | +---+---+      |        (     *  Internet    )
         | +---+---+.     |         ( *                /
         | | vGW  | * * * * * * * * '-'         '-'
         | +-------+ |   | IPsec        \../ \.--/'
         |   +--------+  | Tunnel
         +----------------+
```

          DC Provider Site

          Figure 1 - DC Virtual Network Access via the Internet

. **DC NVO3 virtual network and SP WAN VPN Interconnection**

   In this case, an Enterprise customer wants to use a Service Provider
   (SP) WAN VPN [RFC4364] [RFC7432] to interconnect its sites with an
   NVO3 virtual network in a DC site. The Service Provider constructs a
   VPN for the enterprise customer. Each enterprise site peers with an
   SP PE. The DC Provider and VPN Service Provider can build an NVO3
   virtual network and a WAN VPN independently, and then interconnect
   them via a local link, or a tunnel between the DC GW and WAN
   Provider Edge (PE) devices. The control plane interconnection
   options between the DC and WAN are described in [RFC4364]. Using the
   option A specified in [RFC4364] with VRF-LITE [VRF-LITE], both
   Autonomous System Border Routers (ASBR), i.e., DC GW and SP PE,
   maintain a routing/forwarding table (VRF). Using the option B
   specified in [RFC4364], the DC ASBR and SP ASBR do not maintain the
   VRF table; they only maintain the NVO3 virtual network and VPN
   identifier mappings, i.e., label mapping, and swap the label on the
   packets in the forwarding process. Both option A and B allow the
   NVO3 virtual network and VPN using their own identifiers and two
   identifiers are mapped at DC GW. With the option C in [RFC4364], the
   VN and VPN use the same identifier and both ASBRs perform the tunnel

stitching, i.e., tunnel segment mapping. Each option has pros/cons
[RFC4364] and has been deployed in SP networks depending on the
application requirements. BGP is used in these options for route
distribution between DCs and SP WANs. Note that if the DC is the
SP's Data Center, the DC GW and SP PE in this case can be merged
into one device that performs the interworking of the VN and VPN
within an AS.

These solutions allow the enterprise networks to communicate with
the tenant systems attached to the NVO3 virtual network in the DC
without interfering with the DC provider's underlying physical
networks and other NVO3 virtual networks in the DC. The enterprise
can use its own address space in the NVO3 virtual network. The DC
provider can manage which VM and storage elements attach to the NVO3
virtual network. The enterprise customer manages which applications
run on the VMs without knowing the location of the VMs in the DC.
(See Section 4 for more)

Furthermore, in this use case, the DC operator can move the VMs
assigned to the enterprise from one sever to another in the DC
without the enterprise customer being aware, i.e., with no impact on
the enterprise's 'live' applications. Such advanced technologies
bring DC providers great benefits in offering cloud services, but
add some requirements for NVO3 [RFC7364] as well.

4. DC Applications Using NVO3

NVO3 technology provides DC operators with the flexibility in
designing and deploying different applications in an end-to-end
virtualization overlay environment. The operators no longer need to
worry about the constraints of the DC physical network configuration
when creating VMs and configuring a network to connect them. A DC
provider may use NVO3 in various ways, in conjunction with other
physical networks and/or virtual networks in the DC. This section
highlights some use cases for this goal.

4.1. Supporting Multiple Technologies

Servers deployed in a large data center are often installed at
different times, and may have different capabilities/features. Some
servers may be virtualized, while others may not; some may be
equipped with virtual switches, while others may not. For the
servers equipped with Hypervisor-based virtual switches, some may
support a standardized NVO3 encapsulation, some may not support any
encapsulation, and some may support a documented encapsulation
protocol (e.g. VxLAN [RFC7348], NVGRE [RFC7637]) or proprietary
encapsulations. To construct a tenant network among these servers

and the ToR switches, operators can construct one traditional VLAN
network and two virtual networks where one uses VxLAN encapsulation
and the other uses NVGRE, and interconnect these three networks via
a gateway or virtual GW. The GW performs packet
encapsulation/decapsulation translation between the networks.

Another case is that some software of a tenant has high CPU and
memory consumption, which only makes a sense to run on standalone
servers; other software of the tenant may be good to run on VMs.
However provider DC infrastructure is configured to use NVO3 to
connect VMs and VLAN [IEEE802.1Q] to physical servers. The tenant
network requires interworking between NVO3 and traditional VLAN.

4.2. DC Applications Spanning Multiple Physical Zones

A DC can be partitioned into multiple physical zones, with each zone
having different access permissions and runs different applications.
For example, a three-tier zone design has a front zone (Web tier)
with Web applications, a mid zone (application tier) where service
applications such as credit payment or ticket booking run, and a
back zone (database tier) with Data. External users are only able to
communicate with the Web application in the front zone; the back
zone can only receive traffic from the application zone. In this
case, communications between the zones must pass through one or more
security functions in a physical DMZ zone. Each zone can be
implemented by one NVO3 virtual network and the security functions
in DMZ zone can be used to between two NVO3 virtual networks, i.e.,
two zones. If network functions (NF), especially the security
functions in the physical DMZ can't process encapsulated NVO3
traffic, the NVO3 tunnels have to be terminated for the NF to
perform its processing on the application traffic.

4.3. Virtual Data Center (vDC)

An enterprise data center today may deploy routers, switches, and
network appliance devices to construct its internal network, DMZ,
and external network access; it may have many servers and storage
running various applications. With NVO3 technology, a DC Provider
can construct a virtual Data Center (vDC) over its physical DC
infrastructure and offer a virtual Data Center service to enterprise
customers. A vDC at the DC Provider site provides the same
capability as the physical DC at a customer site. A customer manages
its own applications running in its vDC. A DC Provider can further
offer different network service functions to the customer. The
network service functions may include firewall, DNS, load balancer,
gateway, etc.

Figure 2 below illustrates one such scenario at the service
abstraction level. In this example, the vDC contains several L2 VNs
(L2VNx, L2VNy, L2VNz) to group the tenant systems together on a per-
application basis, and one L3 VN (L3VNa) for the internal routing. A
network firewall and gateway runs on a VM or server that connects to
L3VNa and is used for inbound and outbound traffic processing. A
load balancer (LB) is used in L2VNx. A VPN is also built between the
gateway and enterprise router. An Enterprise customer runs
Web/Mail/Voice applications on VMs within the vDC. The users at the
Enterprise site access the applications running in the vDC via the
VPN; Internet users access these applications via the
gateway/firewall at the provider DC site.

```
                    Internet                    ^ Internet
                                                |
                       ^                   +--+---+
                       |                   |  GW  |
                       |                   +--+---+
                       |                      |
         +-------+--------+             +--+---+
         |Firewall/Gateway+--- VPN-----+router|
         +-------+--------+             +-+--+-+
                 |                       |  |
              ...+....                   |..|
         +-------: L3 VNa :----------+        LANs
      +-+-+      ........           |
      |LB |         |               |      Enterprise Site
      +-+-+         |               |
     ...+...      ...+...         ...+...
     : L2VNx :    : L2VNy :       : L2VNz :
     .......      .......         .......
      |..|         |..|            |..|
      |  |         |  |            |  |
     Web App.    Mail App.      VoIP App.

              Provider DC Site
```
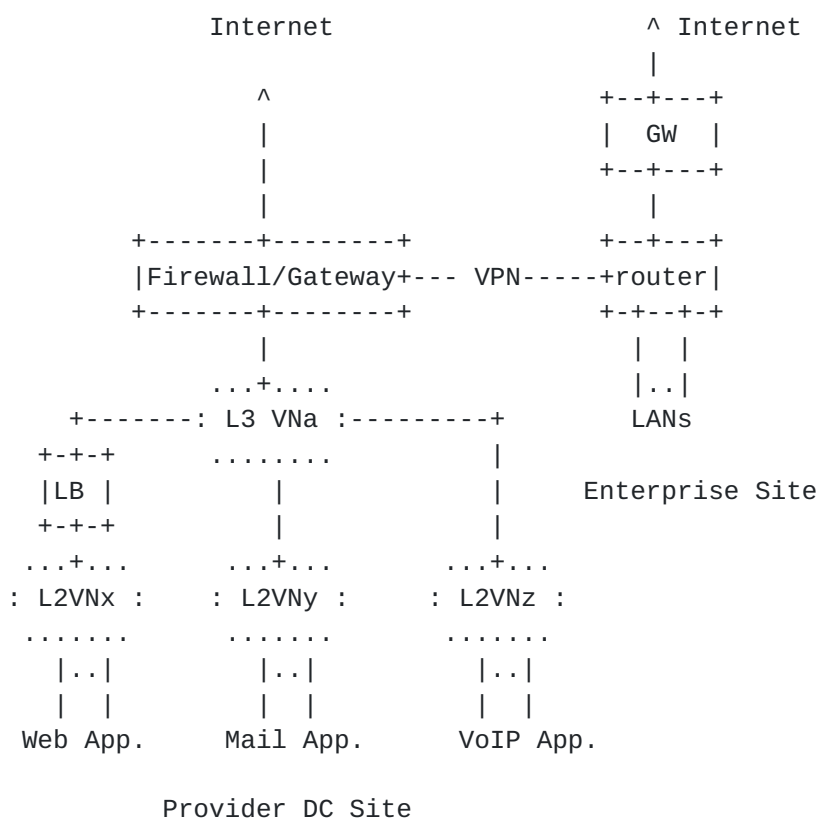
Figure 2 - Virtual Data Center Abstraction View

The enterprise customer decides which applications should be
accessible only via the intranet and which should be assessable via
both the intranet and Internet, and configures the proper security
policy and gateway function at the firewall/gateway. Furthermore, an

enterprise customer may want multi-zones in a vDC (See section 4.2) for the security and/or the ability to set different QoS levels for the different applications.

The vDC use case requires an NVO3 solution to provide DC operators with an easy and quick way to create an NVO3 virtual network and NVEs for any vDC design, to allocate TSs and assign TSs to the corresponding NVO3 virtual network, and to illustrate vDC topology and manage/configure individual elements in the vDC in a secure way.

5. Summary

This document describes some general NVO3 use cases in DCs. The combination of these cases will give operators the flexibility and capability to design more sophisticated support for various cloud applications.

DC services may vary, NVO3 virtual networks make it possible to scale a large number of virtual networks in DC and ensure the network infrastructure not impacted by the number of VMs and dynamic workload changes in DC.

NVO3 uses tunnel techniques to deliver NVO3 traffic over DC physical infrastructure network. A tunnel encapsulation protocol is necessary. An NVO3 tunnel may in turn be tunneled over other intermediate tunnels over the Internet or other WANs.

An NVO3 virtual network in a DC may be accessed by external users in a secure way. Many existing technologies can help achieve this.

6. Security Considerations

Security is a concern. DC operators need to provide a tenant with a secured virtual network, which means one tenant's traffic is isolated from other tenants' traffic and is not leaked to the underlay networks. Tenants are vulnerable to observation and data modification/injection by the operator of the underlay and should only use operators they trust. DC operators also need to prevent a tenant application attacking their underlay DC network; further, they need to protect a tenant application attacking another tenant application via the DC infrastructure network. For example, a tenant application attempts to generate a large volume of traffic to overload the DC's underlying network. This can be done by limiting the bandwidth of such communications.

7. IANA Considerations

   This document does not request any action from IANA.

8. Informative References

   [IEEE802.1Q] IEEE, "IEEE Standard for Local and metropolitan area
            networks -- Media Access Control (MAC) Bridges and Virtual
            Bridged Local Area", IEEE Std 802.1Q, 2011.

   [NIST]   National Institute of Standards and Technology, "The NIST
            Definition of Cloud Computing", SP 880-145, September,
            2011.

   [NVO3MCAST] Ghanwani, A., Dunbar, L., et al, "A Framework for
            Multicast in Network Virtualization Overlays", draft-ietf-
            nvo3-mcast-framework-05, work in progress.

   [RFC1035] Mockapetris, P., "DOMAIN NAMES - Implementation and
            Specification", RFC1035, November 1987.

   [RFC3022] Srisuresh, P. and Egevang, K., "Traditional IP Network
            Address Translator (Traditional NAT)", RFC3022, January
            2001.

   [RFC4301] Kent, S., "Security Architecture for the Internet
            Protocol", rfc4301, December 2005

   [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
            Networks (VPNs)", RFC 4364, February 2006.

   [RFC7348] Mahalingam, M., Dutt, D., et al, "Virtual eXtensible Local
            Area Network (VXLAN): A Framework for Overlaying
            Virtualized Layer 2 Networks over Layer 3 Networks",
            RFC7348 August 2014.

   [RFC7364] Narten, T., et al "Problem Statement: Overlays for Network
            Virtualization", RFC7364, October 2014.

   [RFC7365] Lasserre, M., Motin, T., et al, "Framework for DC Network
            Virtualization", RFC7365, October 2014.

   [RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A. and
            J. Uttaro, "BGP MPLS Based Ethernet VPN", RFC7432,
            February 2015

   [RFC7637] Garg, P., and Wang, Y., "NVGRE: Network Virtualization
             using Generic Routing Encapsulation", RFC7637, Sept. 2015.

   [RFC8014] Black, D., et al, "An Architecture for Overlay Networks
             (NVO3)", rfc8014, January 2017.

   [VRF-LITE] Cisco, "Configuring VRF-lite", http://www.cisco.com

Contributors

     David Black
     Dell EMC
     176 South Street
     Hopkinton, MA 01748
     David.Black@dell.com

     Vinay Bannai
     PayPal
     2211 N. First St,
     San Jose, CA 95131
     Phone: +1-408-967-7784
     Email: vbannai@paypal.com

     Ram Krishnan
     Brocade Communications
     San Jose, CA 95134
     Phone: +1-408-406-7890
     Email: ramk@brocade.com

     Kieran Milne
     Juniper Networks
     1133 Innovation Way
     Sunnyvale, CA 94089
     Phone: +1-408-745-2000
     Email: kmilne@juniper.net

Acknowledgements

 Authors' Addresses

   Lucy Yong
   Huawei Technologies

   Phone: +1-918-808-1918
   Email: lucy.yong@huawei.com

   Linda Dunbar
   Huawei Technologies,
   5340 Legacy Dr.
   Plano, TX 75025 US

   Phone: +1-469-277-5840
   Email: linda.dunbar@huawei.com

   Mehmet Toy
   Verizon

   E-mail : mtoy054@yahoo.com

   Aldrin Isaac
   Juniper Networks
   E-mail: aldrin.isaac@gmail.com

   Vishwas Manral

   Email: vishwas@ionosnetworks.com