L2VPN Workgroup   Ali Sajassi
INTERNET-DRAFT   Samer Salam
Intended Status: Standards Track   Cisco

                                                              Yakov Rekhter
Wim Henderickx   John Drake
Alcatel-Lucent   Juniper

                                                                 Lucy Yong
Florin Balus   Linda Dunbar
Nuage Networks   Huawei

Expires: January 15, 2014   July 15, 2013

IP Inter-Subnet Forwarding in EVPN
draft-sajassi-l2vpn-evpn-inter-subnet-forwarding-02

Abstract

EVPN provides an extensible and flexible multi-homing VPN solution for intra-subnet connectivity among hosts/VMs over an MPLS/IP network. However, there are scenarios in which inter-subnet forwarding among hosts/VMs across different IP subnets is required, while maintaining the multi-homing capabilities of EVPN. This document describes an IRB solution based on EVPN to address such requirements.

Status of this Memo

Table of Contents

Terminology

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].

IRB: Integrated Routing and Bridging

IRB Interface: A virtual interface that connects the bridging module and the routing module on an NVE.

NVE: Network Virtualization Endpoint

1 Introduction

EVPN provides an extensible and flexible multi-homing VPN solution for intra-subnet connectivity among hosts/VMs over an MPLS/IP network. However, there are scenarios where, in addition to intra-subnet forwarding, inter-subnet forwarding is required among hosts/VMs across different IP subnets at the EVPN PE nodes, also known as EVPN NVE nodes throughout this document, while maintaining the multi-homing capabilities of EVPN. This document describes an IRB solution based on EVPN to address such requirements.

1.1 Traditional Inter-Subnet Forwarding

The inter-subnet communication is traditionally achieved at the L3 Gateway nodes where all the inter-subnet communication policies are enforced. Even for different subnets belonging to one IP-VPN or tenant, traffic may need to go through FW or IPS between the trusted and un-trusted zones.

Some operators may prefer centralized approach, i.e. only have a set of default L3 gateways (whose redundancy is typically achieved by VRRP) for all inter-subnet traffic to go through. Usually there are FW, IPS, or other network appliances directly attached to the centralized L3 Gateway nodes. The centralized approach makes it easier for maintaining consistent policies and less prone to configuration errors. However, such centralized approach suffers from a major drawback of requiring all traffic to be hair-pinned to the L3GW nodes.

Some operators may prefer fully distributed L3 gateway design, e.g. allowing all NVEs to have the policies to route traffic across subnets. Under this design, all traffic between hosts attached to one NVE can be routed locally, thus avoiding traffic hair-pinning issue at the centralized L3GW. The perceived drawback of this fully distributed approach may be the extra effort required in maintaining policy consistence across all the NVEs.

Some operators may prefer somewhere in the middle, i.e. allowing NVEs to route traffic across only selected subnets. For example, allow NVEs to route traffic among subnets belonging to one tenant or one security zone.

1.2. Scenarios of EVPN NVEs as L3GW

When an EVPN NVE node is not the L3GW for the subnets attached, the EVPN NVE performs only L2 switching function for the traffic initiated from or destined to the hosts attached to the NVE.

Some EVPN NVEs can be the default L3GWs for some subnets. In this situation, the EVPN NVEs can route traffic across the subnets for which they are default L3GWs.

When there are multiple subnets attached to an EVPN NVE, some of the subnets could have the EVPN NVE as their L3GW, some other subnets don't have the NVE as their L3GW. For example: "Subnet-X" can communicate with "Subnet-Y" via NVE "A", but "Subnet-X" can't communicate with "Subnet-Z" via NVE "A". So when the "Subnet-X" needs to communicate with "Subnet-Z", the traffic might need to be routed through another device (e.g. FW, IPS, or another L3GW node).

1. When the EVPN NVE is the L3GW for "Subnet -X", hosts within "Subnet-X" will have the NVE's IRB MAC address as their default GW MAC address when they send data frames towards targets in different subnets.

2. When the EVPN NVE is not the L3GW for "Subnet-Y", hosts within "Subnet-Y", (even though still attached to the NVE), will use their own designated L3GW MAC address (that is different from the NVE's IRB address) in data frames destined towards targets in different subnets.

2 Inter-Subnet Forwarding Scenarios

The inter-subnet forwarding scenarios performed by an EVPN NVE can be divided into the following five categories. The last scenario, along with their corresponding solutions, are described in [EVPN-IPVPN-INTEROP]. The solutions for the first four scenarios are the focus of this document.

1. Switching among EVPN instances within a DC

2. Switching among EVPN instances in different DCs without route aggregation

3. Switching among EVPN instances in different DCs with route aggregation

4. Switching among IP-VPN sites and EVPN instances with route aggregation

5. Switching among IP-VPN sites and EVPN instances without route aggregation

In the above scenario, the term "route aggregation" refers to the

case where for a given EVI/VRF a node situated at the WAN edge of the data center network behaves as a default gateway for all the destinations that are outside the data center. The absence of route aggregation refers to the scenario where a given EVI/VRF within a data center has (host) routes to individual VMs that are outside of the data center.

In the case (4) the WAN edge node also performs route aggregation for all the destinations within its own data center, and acts as an interworking unit between EVPN and IP VPN (it implements both EVPN and IP VPN functionality).

```
                             +---+    Enterprise Site 1
                             |PE1|----- H1
                             +---+
                               /
                         ,---------.             Enterprise Site 2
                       ,'           `.    +---+
        ,---------.  /(    MPLS/IP    )---|PE2|-----  H2
       '   DCN 3   `./ `.   Core    ,'    +---+
        `-+------+'     `-+------+'
        __/__           / /      \ \
       :NVE4 :        +---+       \ \
       '-----'   ,----|GW |.       \ \
          |    ,'     +---+ `.      ,---------.
         VM6  (      DCN 1    )   ,'           `.
               `.           ,'   (      DCN 2    )
                 `-+------+'      `.           ,'
                   __/__            `-+------+'
                  :NVE1 :           __/__   __\__
                  '-----'          :NVE2 :  :NVE3 :
                   |  |            '-----'  '-----'
                  VM1 VM2            |  |      |
                                    VM3 VM4   VM5
```

Figure 2: Interoperability Use-Cases

In what follows, we will describe scenarios 3 through 6 in more detail.

2.1 Switching among EVIs within a DC

In this scenario, connectivity is required between hosts (e.g. VMs) in the same data center, where those hosts belong to different IP subnets. All these subnets are part of the same IP VPN. Each subnet is associated with a single EVPN, where each such EVPN is realized by a collection of EVIs residing on appropriate NVEs.

As an example, consider VM3 and VM5 of Figure 2 above. Assume that connectivity is required between these two VMs where VM3 belongs to the IP3 subnet whereas VM5 belongs to the IP5 subnet. Both IP3 and IP5 subnets are part of the same IP VPN. NVE2 has an EVI3 associated with IP3 subnet and NVE3 has an EVI5 associated with the IP5 subnet.

2.2 Switching among EVIs in different DCs without route aggregation

This case is similar to that of section 2.1 above albeit for the fact that the hosts belong to different data centers that are interconnected over a WAN (e.g. MPLS/IP PSN). The data centers in question here are seamlessly interconnected to the WAN, i.e., the WAN edge does not maintain any host/VM-specific addresses in the forwarding path.

As an example, consider VM3 and VM6 of Figure 2 above. Assume that connectivity is required between these two VMs where VM3 belongs to the IP3 subnet whereas VM6 belongs to the IP6 subnet. NVE2 has an EVI3 associated with IP3 subnet and NVE4 has an EVI6 associated with the IP6 subnet. Both IP3 and IP6 subnets are part of the same IP VPN and both EVI3 and EVI6 are associated with their VRFs for that IP VPN.

2.3 Switching among EVIs in different DCs with route aggregation

In this scenario, connectivity is required between hosts (e.g. VMs) in different data centers, and those hosts belong to different IP subnets. What makes this case different from that of Section 2.2 is that (in the context of a given EVI/VRF) at least one of the data centers in question has a gateway as the WAN edge switch. Because of that, the EVIs/VRFs within each data center need not maintain (host) routes to individual VMs outside of the data center.

As an example, consider VM1 and VM5 of Figure 2 above. Assume that connectivity is required between these two VMs where VM1 belongs to the IP1 subnet whereas VM5 belongs to the IP5 subnet thus IP1 and IP5 subnets belong to the same IP VPN. NVE3 has an EVI5 associated with the IP5 subnet and NVE1 has an EVI1 associated with the IP1 subnet. Both EVI1 and EVI5 have associated with their VRFs that belong to the IP VPN that includes IP1 and IP5 subnets. Due to the gateway at the edge of DCN 1, NVE1 does not have the address of VM5 in its VRF table but instead it has a default route in its VRF with the next-hop being the GW.

2.4 Switching among IP-VPN sites and EVIs with route aggregation

In this scenario (within a context of a particular EVPN instance), connectivity is required between hosts (e.g. VMs) in a data center and hosts in an enterprise site that belongs to a given IP-VPN. The NVE within the data center is an EVPN NVE, whereas the enterprise site has an IP-VPN PE. Furthermore, the data center in question has a gateway as the WAN edge switch. Because of that, the NVE in the data center does not need to maintain individual IP prefixes advertised by enterprise sites (by IP-VPN PEs).

As an example, consider end-station H1 and VM2 of Figure 2. Assume that connectivity is required between the end-station and the VM, where VM2 belongs to the IP2 subnet that is realized using EVPN, whereas H1 belongs to an IP VPN site connected to PE1 (PE1 maintains an IP VPN VRF associated with that IP VPN). NVE1 has an EVI2 associated with the IP2 subnet. Moreover, NVE1 maintains a VRF associated with EVI2. PE1 originates a VPN-IP route that covers H1. The gateway at the edge of DCN1 performs interworking function between IP-VPN and EVPN. As a result of this, a default route in the VRF associated with EVI2, pointing to the gateway as the next hop, and a route to the VM2 (or maybe IP2 subnet) on the H1's VRF on PE1 are sufficient for the connectivity between H1 and VM2.

3 Default L3 Gateway Addressing

3.1 Homogeneous Environment

This is an environment where all NVEs to which an EVPN instance could potentially be attached (or moved), perform inter-subnet switching. Therefore, inter-subnet traffic can be locally switched by the EVPN NVE connecting the VMs belonging to different subnets.

To support such inter-subnet forwarding, the NVE behaves as an IP Default Gateway from the perspective of the attached end-stations (e.g. VMs). Two models are possible, as discussed in [DC-MOBILITY]:

1. All the EVIs of a given EVPN instance use the same anycast default gateway IP address and the same anycast default gateway MAC address. On each NVE, this default gateway IP/MAC address correspond to the IRB interface of the EVI associated with that EVPN instance.

2. Each EVI of a given EVPN instance uses its own default gateway IP and MAC addresses, and these addresses are aliased to the same conceptual gateway through the use of the Default Gateway extended community as specified in [EVPN], which is carried in the EVPN MAC Advertisement routes. On each NVE, this default gateway IP/MAC address correspond to the IRB interface of the EVI associated with that EVPN instance.

Both of these models enable a packet forwarding paradigm where inter-subnet traffic can bypass the VRF processing on the egress (i.e. disposition) NVE. The egress NVE merely needs to perform a lookup in the associated EVI and forward the Ethernet frames unmodified, i.e. without rewriting the source MAC address. This is different from traditional IRB forwarding where a packet is forwarded through the bridge module followed by the routing module on the ingress NVE, and then forwarded through the routing module followed by the bridging module on the egress NVE. For inter-subnet forwarding using EVPN, the routing module on the egress NVE can be completely bypassed.

It is worth noting that if the applications that are running on the hosts (e.g. VMs) are employing or relying on any form of MAC security, then the first model (i.e. using anycast addresses) would be required to ensure that the applications receive traffic from the same source MAC address that they are sending to.

3.1 Heterogeneous Environment

For large data centers with thousands of servers and ToR (or Access) switches, some of them may not have the capability of maintaining or enforcing policies for inter-subnet switching. Even though policies among multiple subnets belonging to same tenant can be simpler, hosts belonging to one tenant can also send traffic to peers belonging to different tenants or security zones. A L3GW not only needs to enforce policies for communication among subnets belonging to a single tenant, but also it needs to know how to handle traffic destined towards peers in different tenants. Therefore, there can be a mixed environment where an NVE performs inter-subnet switching for some EVPN instances but not others.

4 Operational Models for Inter-Subnet Forwarding

4.1 Among EVPN NVEs within a DC

When an EVPN MAC advertisement route is received by the NVE, the IP address associated with the route is used to populate the VRF, whereas the MAC address associated with the route is used to populate both the bridge-domain MAC table, as well as the adjacency associated with the IP route in the VRF.

When an Ethernet frame is received by an ingress NVE, it performs a lookup on the destination MAC address in the associated EVI. If the MAC address corresponds to its IRB Interface MAC address, the ingress NVE deduces that the packet MUST be inter-subnet routed. Hence, the ingress NVE performs an IP lookup in the associated VRF table. The lookup identifies both the next-hop (i.e. egress) NVE to which the

packet must be forwarded, in addition to an adjacency that contains a MAC rewrite and an MPLS label stack. The MAC rewrite holds the MAC address associated with the destination host (as populated by the EVPN MAC route), instead of the MAC address of the next-hop NVE. The ingress NVE then rewrites the destination MAC address in the packet with the address specified in the adjacency. It also rewrites the source MAC address with its IRB Interface MAC address. The ingress NVE, then, forwards the frame to the next-hop (i.e. egress) NVE after encapsulating it with the MPLS label stack. Note that this label stack includes the LSP label as well as the EVI label that was advertised by the egress NVE. When the MPLS encapsulated packet is received by the egress NVE, it uses the EVI label to identify the bridge-domain table. It then performs a MAC lookup in that table, which yields the outbound interface to which the Ethernet frame must be forwarded. Figure 2 below depicts the packet flow, where NVE1 and NVE2 are the ingress and egress NVEs, respectively.

```
                   NVE1                NVE2
             +------------+     +------------+
             | ...   ...  |     | ...   ...  |
             |(EVI)-(VRF) |     |(VRF)-(EVI) |
             | .|.   .|.  |     | ...   |..| |
             +------------+     +------------+
                ^     v                 ^  V
                |     |                 |  |
          VM1->-+     +-->--------------+  +->-VM2
```

Figure 2: Inter-Subnet Forwarding Among EVPN NVEs within a DC

Note that the forwarding behavior on the egress NVE is similar to EVPN intra-subnet forwarding. In other words, all the packet processing associated with the inter-subnet forwarding semantics is confined to the ingress NVE.

It should also be noted that [EVPN] provides different level of granularity for the EVI label. Besides identifying bridge domain table, it can be used to identify the egress interface or a destination MAC address on that interface. If EVI label is used for egress interface or destination MAC address identification, then no MAC lookup is needed in the egress EVI and the packet can be directly forwarded to the egress interface just based on EVI label lookup.

4.2 Among EVPN NVEs in Different DCs Without Route Aggregation

When an EVPN MAC advertisement route is received by the NVE, the IP address associated with the route is used to populate the VRF,

whereas the MAC address associated with the route is used to populate both the bridge-domain MAC table, as well as the adjacency associated with the IP route in the VRF.

When an Ethernet frame is received by an ingress NVE, it performs a lookup on the destination MAC address in the associated EVI. If the MAC address corresponds to its IRB Interface MAC address, the ingress NVE deduces that the packet MUST be inter-subnet routed. Hence, the ingress NVE performs an IP lookup in the associated VRF table. The lookup identifies both the next-hop (i.e. egress) Gateway to which the packet must be forwarded, in addition to an adjacency that contains a MAC rewrite and an MPLS label stack. The MAC rewrite holds the MAC address associated with the destination host (as populated by the EVPN MAC route), instead of the MAC address of the next-hop Gateway. The ingress NVE then rewrites the destination MAC address in the packet with the address specified in the adjacency. It also rewrites the source MAC address with its IRB Interface MAC address. The ingress NVE, then, forwards the frame to the next-hop (i.e. egress) Gateway after encapsulating it with the MPLS label stack. Note that this label stack includes the LSP label as well as an EVI label. The EVI label could be either advertised by the ingress Gateway, if inter-AS option B is used, or advertised by the egress NVE, if inter-AS option C is used. When the MPLS encapsulated packet is received by the ingress Gateway, the processing again differs depending on whether inter-AS option B or option C is employed: in the former case, the ingress Gateway swaps the EVI label in the packets with the EVI label value received from the egress Gateway. In the latter case, the ingress Gateway does not modify the EVI label and performs normal label switching on the LSP label. Similarly on the egress Gateway, for option B, the egress Gateway swaps the EVI label with the value advertised by the egress NVE. Whereas, for option C, the egress Gateway does not modify the EVI label, and performs normal label switching on the LSP label. When the MPLS encapsulated packet is received by the egress NVE, it uses the EVI label to identify the bridge-domain table. It then performs a MAC lookup in that table, which yields the outbound interface to which the Ethernet frame must be forwarded. Figure 3 below depicts the packet flow.

```
            NVE1            GW1             GW2            NVE2
      +------------+  +------------+  +------------+  +------------+
      | ...   ...  |  |    ...     |  |    ...     |  | ...   ...  |
      |(EVI)-(VRF) |  |   [LS ]    |  |   [LS ]    |  |(VRF)-(EVI) |
      | .|.   .|.  |  |    |..|    |  |    |..|    |  | ...   |..| |
      +------------+  +------------+  +------------+  +------------+
         ^     v           ^  V            ^  V                ^  V
         |     |           |  |            |  |                |  |
   VM1->-+     +-->--------+  +------------+  +----------------+  +->-VM2
```

Figure 3: Inter-Subnet Forwarding Among EVPN NVEs in Different DCs without Route Aggregation

4.3 Among EVPN NVEs in Different DCs with Route Aggregation

In this scenario, the NVEs within a given data center do not have entries for the MAC/IP addresses of hosts in remote data centers. Rather, the NVEs have a default IP route pointing to the WAN gateway for each VRF. This is accomplished by the WAN gateway advertising for a given EVPN that spans multiple DC a default VPN-IP route that is imported by the NVEs of that EVPN that are in the gateway's own DC.

When an Ethernet frame is received by an ingress NVE, it performs a lookup on the destination MAC address in the associated EVI. If the MAC address corresponds to the IRB Interface MAC address, the ingress NVE deduces that the packet MUST be inter-subnet routed. Hence, the ingress NVE performs an IP lookup in the associated VRF table. The lookup, in this case, matches the default route which points to the local WAN gateway. The ingress NVE then rewrites the destination MAC address in the packet with the IRB Interface MAC address of the local WAN gateway. It also rewrites the source MAC address with its own IRB Interface MAC address. The ingress NVE, then, forwards the frame to the WAN gateway after encapsulating it with the MPLS label stack. Note that this label stack includes the LSP label as well as the IP-VPN label that was advertised by the local WAN gateway. When the MPLS encapsulated packet is received by the local WAN gateway, it uses the IP-VPN label to identify the VRF table. It then performs an IP lookup in that table. The lookup identifies both the remote WAN gateway (of the remote data center) to which the packet must be forwarded, in addition to an adjacency that contains a MAC rewrite and an MPLS label stack. The MAC rewrite holds the MAC address associated with the ultimate destination host (as populated by the EVPN MAC route). The local WAN gateway then rewrites the destination MAC address in the packet with the address specified in the adjacency. It also rewrites the source MAC address with its IRB Interface MAC address. The local WAN gateway, then, forwards the frame to the remote WAN gateway after encapsulating it with the MPLS label stack. Note that

this label stack includes the LSP label as well as a VPN label that was advertised by the remote WAN gateway. When the MPLS encapsulated packet is received by the remote WAN gateway, it simply swaps the VPN label with the EVI label advertised by the egress NVE. This implies that the remote WAN gateway must allocate the VPN label at least at the granularity of a (VRF, egress NVE) tuple. The remote WAN gateway then forward the packet to the egress NVE. The egress NVE then performs a MAC lookup in the EVI (identified by the received EVI label) to determine the outbound port to send the traffic on.

Figure 4 below depicts the forwarding model.

```
            NVE1            GW1             GW2            NVE2
      +------------+  +------------+  +------------+  +------------+
      | ...   ...  |  | ...   ...  |  |    ...     |  | ...   ...  |
      |(EVI)-(VRF) |  |(VRF)-(EVI) |  |   [LS ]    |  |(VRF)-(EVI) |
      | .|.   .|.  |  | |..|       |  |   |...|    |  | ...   |..| |
      +------------+  +------------+  +------------+  +------------+
         ^     v         ^  V              ^   V               ^  V
         |     |         |  |              |   |               |  |
   VM1->-+     +-->-----+  +--------------+   +---------------+  +->-VM2
```

Figure 4: Inter-Subnet Forwarding Among EVPN NVEs in Different DCs with Route Aggregation

4.4 Among IP-VPN Sites and EVPN NVEs with Route Aggregation

In this scenario, the NVEs within a given data center do not have entries for the IP addresses of hosts in remote enterprise sites. Rather, the NVEs have a default IP route pointing to the WAN gateway for each VRF.

When an Ethernet frame is received by an ingress NVE, it performs a lookup on the destination MAC address in the associated EVI. If the MAC address corresponds to the IRB Interface MAC address, the ingress NVE deduces that the packet MUST be inter-subnet routed. Hence, the ingress NVE performs an IP lookup in the associated VRF table. The lookup, in this case, matches the default route which points to the local WAN gateway. The ingress NVE then rewrites the destination MAC address in the packet with the IRB Interface MAC address of the local WAN gateway. It also rewrites the source MAC address with its own IRB Interface MAC address. The ingress NVE, then, forwards the frame to the WAN gateway after encapsulating it with the MPLS label stack. Note that this label stack includes the LSP label as well as the IP-VPN label that was advertised by the local WAN gateway. When the MPLS encapsulated packet is received by the local WAN gateway, it uses the

IP-VPN label to identify the VRF table. It then performs an IP lookup in that table. The lookup identifies the next hop ASBR to which the packet must be forwarded. The local gateway in this case strips the Ethernet encapsulation and forwards the IP packet to the ASBR using a label stack comprising of an LSP label and a VPN label that was advertised by the ASBR. When the MPLS encapsulated packet is received by the ASBR, it simply swaps the VPN label with the IP-VPN label advertised by the egress PE. This implies that the remote WAN gateway must allocate the VPN label at least at the granularity of a (VRF, egress PE) tuple. The ASBR then forwards the packet to the egress PE. The egress PE then performs an IP lookup in the VRF (identified by the received IP-VPN label) to determine where to forward the traffic.

Figure 5 below depicts the forwarding model.

```
         NVE1            GW1             ASBR          NVE2
   +------------+  +------------+  +------------+  +------------+
   | ...   ...  |  | ...   ...  |  |    ...     |  |        ... |
   |(EVI)-(VRF) |  |(VRF)-(EVI) |  |   [LS ]    |  |       (VRF)|
   | .|.   .|.  |  | |..|       |  |   |...|    |  |       |..| |
   +------------+  +------------+  +------------+  +------------+
      ^     v         ^  V              ^   V               ^  V
      |     |         |  |              |   |               |  |
VM1->-+     +-->-----+  +--------------+   +---------------+  +->-H1
```

Figure 5: Inter-Subnet Forwarding Among IP-VPN Sites and EVPN NVEs with Route Aggregation

4.5 Use of Centralized Gateway

In this scenario, the NVEs within a given data center need to forward traffic in L2 to a centralized L3GW for a number of reasons: a) they don't have IRB capabilities or b) they don't have required policy for switching traffic between different tenants or security zones. The centralized L3GW performs both the IRB function for switching traffic among different EVPN instances as well as it performs interworking function when the traffic needs to be switched between IP-VPN sites and EVPN instances.

5 VM Mobility

5.1 VM Mobility & Optimum Forwarding for VM's Outbound Traffic

Optimum forwarding for the VM's outbound traffic, upon VM mobility, can be achieved using either the anycast default Gateway MAC and IP

addresses, or using the address aliasing as discussed in [DC-MOBILITY].

5.2 VM Mobility & Optimum Forwarding for VM's Inbound Traffic

For optimum forwarding of the VM's inbound traffic, upon VM mobility, all the NVEs and/or IP-VPN PEs need to know the up to date location of the VM. Two scenarios must be considered, as discussed next.

In what follows, we use the following terminology:

- source NVE refers to the NVE behind which the VM used to reside prior to the VM mobility event.

- target NVE refers to the new NVE behind which the VM has moved after the mobility event.

5.2.1 Mobility without Route Aggregation

In this scenario, when a target NVE detects that a MAC mobility event has occurred, it initiates the MAC mobility handshake in BGP as specified in [EVPN]. The WAN Gateways, acting as ASBRs in this case, re-advertise the MAC route of the target NVE with the MAC Mobility extended community attribute unmodified. Because the WAN Gateway for a given data center re-advertises BGP routes received from the WAN into the data center, the source NVE will receive the MAC Advertisement route of the target NVE (with the next hop attribute adjusted depending on which inter-AS option is employed). The source NVE will then withdraw its original MAC Advertisement route as a result of evaluating the Sequence Number field of the MAC Mobility extended community in the received MAC Advertisement route. This is per the procedures already defined in [EVPN].

5.2.2 Mobility with Route Aggregation

This section will be completed in the next revision.

6 Acknowledgements

The authors would like to thank Sami Boutros for his valuable comments.

7 Security Considerations

8 IANA Considerations

9 References

9.1 Normative References

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

9.2 Informative References

[EVPN] Sajassi et al., "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-04.txt, work in progress, July, 2014.

[EVPN-IPVPN-INTEROP] Sajassi et al., "EVPN Seamless Interoperability with IP-VPN", draft-sajassi-l2vpn-evpn-ipvpn-interop-01, work in progress, October, 2012.

[DC-MOBILITY] Aggarwal et al., "Data Center Mobility based on BGP/MPLS, IP Routing and NHRP", draft-raggarwa-data-center-mobility-05.txt, work in progress, June, 2013.

Authors' Addresses

Ali Sajassi
Cisco
Email: sajassi@cisco.com

Samer Salam
Cisco
Email: ssalam@cisco.com

Yakov Rekhter
Juniper Networks
Email: yakov@juniper.net

John E. Drake
Juniper Networks
Email: jdrake@juniper.net

Lucy Yong
Huawei Technologies
Email: lucy.yong@huawei.com

Linda Dunbar
Huawei Technologies
Email: linda.dunbar@huawei.com

Wim Henderickx
Alcatel-Lucent
Email: wim.henderickx@alcatel-lucent.com

Florin Balus
Alcatel-Lucent
Email: Florin.Balus@alcatel-lucent.com