

INTERNET-DRAFT

N. Malhotra

A. Sajassi

A. Pattekar

Intended Status: Proposed Standard

(Cisco)

A. Lingala

(AT&T)

Expires: December 29, 2017

June 27, 2017

**Extended Mobility Procedures for EVPN-IRB
draft-malhotra-bess-evpn-irb-extended-mobility-00**

Abstract

Procedure to handle host mobility in a layer 2 Network with EVPN control plane is defined as part of [RFC 7432](#). EVPN has since evolved to find wider applicability across various IRB use cases that include distributing both MAC and IP reachability via a common EVPN control plane. MAC Mobility procedures defined in [RFC 7432](#) are extensible to IRB use cases if a fixed 1:1 mapping between VM IP and MAC is assumed across VM moves. Generic mobility support for IP and MAC that allows these bindings to change across moves is required to support a broader set of EVPN IRB use cases, and requires further consideration. EVPN-LAG based multi-homing further introduces scenarios that require additional consideration from mobility perspective. Intent of this draft is to enumerate a set of design considerations applicable to mobility across EVPN IRB use cases and define generic sequence number assignment procedures to address these IRB use cases.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at

<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

Copyright and License 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 [4](#)
- [1.1](#) Terminology [5](#)
- [2](#). Optional MAC only RT-2 [5](#)
- [3](#). Mobility Use Cases [5](#)
- [3.1](#) VM MAC+IP Move [5](#)
- [3.2](#) VM IP Move to new MAC [6](#)
- [3.2.1](#) VM Reload [6](#)
- [3.2.2](#) MAC Sharing [6](#)
- [3.2.3](#) Problem [6](#)
- [3.3](#) VM MAC move to new IP [7](#)
- [3.3.1](#) Problem [7](#)
- [4](#). Multi-homed hosts via ELAG [9](#)
- [5](#). Design Considerations [10](#)
- [6](#). Solution Components [11](#)
- [6.1](#) Sequence Number Inheritance [11](#)
- [6.2](#) MAC Sharing [12](#)
- [6.3](#) Multi-homing Mobility Synchronization [13](#)
- [7](#). Requirements for Sequence Number Assignment [13](#)
- [7.1](#) LOCAL MAC-IP learning [13](#)
- [7.2](#) LOCAL MAC learning [14](#)
- [7.3](#) Remote MAC OR MAC-IP Update [14](#)
- [7.4](#) REMOTE (SYNC) MAC update [14](#)
- [7.5](#) REMOTE (SYNC) MAC-IP update [15](#)

- [7.6 Inter-op](#) [15](#)
- [8. Duplicate Host Detection](#) [15](#)
 - [8.1 Duplicate IP detection Procedure](#) [16](#)
 - [8.2 Duplicate Host Recovery](#) [17](#)
 - [8.2.1 Route Un-freezing CLI](#) [17](#)
 - [8.2.2 Route clearing CLI](#) [17](#)
- [9. Security Considerations](#) [18](#)
- [10. IANA Considerations](#) [18](#)
- [11. References](#) [18](#)
 - [11.1 Normative References](#) [18](#)
 - [11.2 Informative References](#) [18](#)
- [12. Acknowledgements](#) [18](#)
- Authors' Addresses [18](#)
- [Appendix A](#) [18](#)

1 Introduction

EVPN-IRB enables capability to advertise both MAC and IP routes via a single MAC+IP RT-2 advertisement. MAC is imported into local bridge MAC table and enables L2 bridged traffic across the network overlay. IP is imported into the local ARP table in an asymmetric IRB design OR imported into the IP routing table in a symmetric IRB design, and enables routed traffic across the layer 2 network overlay. To support EVPN mobility procedure, a single sequence number mobility attribute is advertised with the combined MAC+IP route. A single sequence number advertised with the combined MAC+IP route to resolve both MAC and IP reachability implicitly assumes a 1:1 fixed mapping between IP and MAC.

While a fixed 1:1 mapping between IP and MAC is a common use case that could be addressed via existing MAC mobility procedure, additional IRB scenarios need to be considered, that don't necessarily adhere to this assumption. Following IRB mobility scenarios are considered:

- o VM move results in VM IP and MAC moving together
- o VM move results in VM IP moving to a new MAC association
- o VM move results in VM MAC moving to a new IP association

While existing MAC mobility procedure can be leveraged for MAC+IP move in the first scenario, subsequent scenarios result in a new MAC-IP association. As a result, a single sequence number assigned independently per-[MAC, IP] is not sufficient to determine most recent reachability for both MAC and IP, unless the sequence number assignment algorithm is designed to allow for changing MAC-IP bindings across moves.

Purpose of this draft is to define additional sequence number assignment and handling procedures to adequately address generic mobility support across EVPN-IRB overlay use cases that allow MAC-IP bindings to change across VM moves and can support mobility for both MAC and IP components carried in an EVPN RT-2 for these use cases.

In addition, for hosts on an ESI multi-homed to multiple GW devices, additional procedure is proposed to ensure synchronized sequence number assignments across the multi-homing devices.

Content presented in this draft is independent of data plane encapsulation used in the overlay being MPLS or VXLAN. It is also largely independent of the EVPN IRB solution being based on symmetric OR asymmetric IRB design. In other words, ideas presented in this

draft should apply equally to symmetric and asymmetric EVPN IRB designs, as well as MPLS and VXLAN based overlays.

1.1 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](#)].

ARP is widely referred to in this document. This is simply for ease of reading, and as such, these references are equally applicable to ND (neighbor discovery) as well.

Term GW used widely in the document refers to an IRB GW that is doing routing and bridging between an access network and an EVPN enabled overlay network.

2. Optional MAC only RT-2

In an EVPN IRB scenario, where a single MAC+IP RT-2 advertisement carries both IP and MAC routes, a MAC only RT-2 advertisement is redundant for host MACs that are advertised via MAC+IP RT-2. As a result, a MAC only RT-2 is an optional route that may not be advertised from or received at an IRB GW. This is an important consideration for mobility scenarios discussed in subsequent sections.

MAC only RT-2 may still be advertised for non-IP host MACs that are not advertised via MAC+IP RT-2.

3. Mobility Use Cases

This section goes over IRB mobility use cases taken into consideration while defining mobility procedures proposed in later sections:

- o VM move results in VM IP and MAC moving together
- o VM move results in VM IP moving to a new MAC association
- o VM move results in VM MAC moving to a new IP association

3.1 VM MAC+IP Move

This is the baseline case, wherein a VM move results in both VM MAC and IP moving together with no change in MAC-IP binding across a move. Existing MAC mobility defined in [RFC 7432](#) may be extrapolated

to apply to corresponding MAC+IP route to support this mobility scenario.

3.2 VM IP Move to new MAC

This is the case, where a VM move results in VM IP moving to a new MAC binding.

3.2.1 VM Reload

A VM reload or an orchestrated VM move that results in VM being re-spawned at a new location may result in VM getting a new MAC assignment, while maintaining existing IP address. This results in a VM IP move to a new MAC binding:

IP-a, MAC-a ---> IP-a, MAC-b

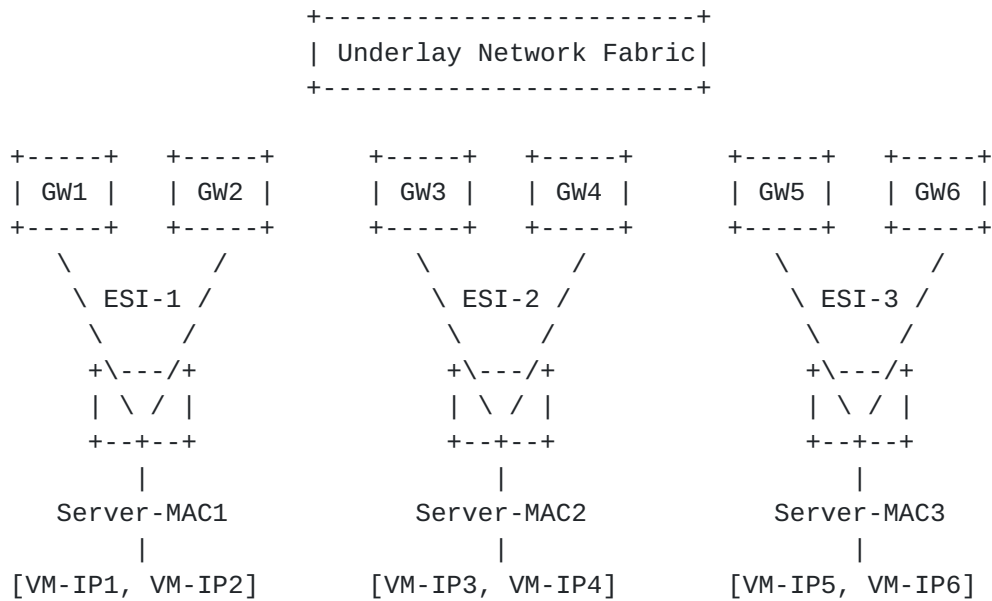
3.2.2 MAC Sharing

This takes into account scenarios, where multiple hosts, each with a unique IP, may share a common MAC binding, and a host move results in a new MAC binding for the host IP.

As an example, host VMs running on a single physical server, each with a unique IP, may share the same physical server MAC. In yet another scenario, an L2 access network may be behind a firewall, such that all hosts IPs on the access network are learnt with a common firewall MAC. In all such "shared MAC" use cases, multiple local MAC-IP ARP entries may be learnt with the same MAC. A VM IP move, in such scenarios, could result in new MAC association for the VM IP.

3.2.3 Problem

In both of the above scenarios, a combined MAC+IP EVPN RT-2 advertised with a single sequence number attribute implicitly assumes a fixed IP to MAC mapping. A host IP move to a new MAC breaks this assumption and results in a new MAC+IP route. If this new MAC+IP route is independently assigned a new sequence number, the sequence number can no longer be used to determine most recent host IP reachability in a symmetric EVPN-IRB design OR the most recent IP to MAC binding in an asymmetric EVPN-IRB design.



As an example, consider the above topology with host VMs sharing the physical server MAC. In steady state, [IP1, MAC1] route is learnt at [GW1, GW2] and advertised to remote GWs with a sequence number N. Now, VM-IP1 is moved to Server-MAC2. ARP or ND based local learning at [GW3, GW4] would now result in a new [IP1, MAC2] route being learnt. If route [IP1, MAC2] is learnt as a new MAC+IP route and assigned a new sequence number of say 0, mobility procedure for VM-IP1 will not trigger across the overlay network.

A clear sequence number assignment procedure needs to be defined to unambiguously determine the most recent IP reachability, IP to MAC binding, and MAC reachability for such a MAC sharing scenario.

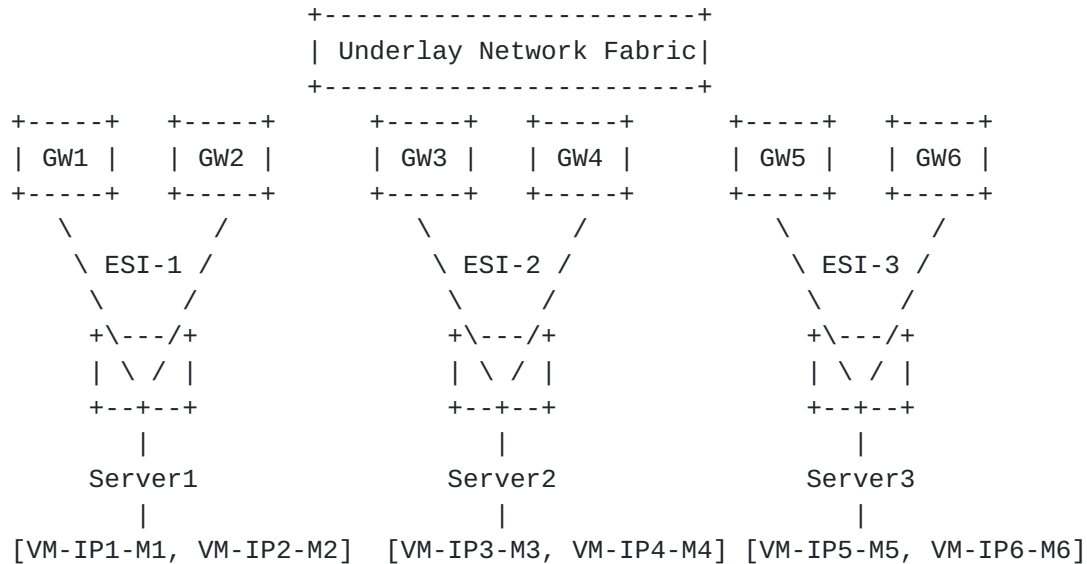
3.3 VM MAC move to new IP

This is a scenario where host move or re-provisioning behind a new gateway location may result in the same VM MAC getting a new IP address assigned.

3.3.1 Problem

Complication with this scenario is that MAC reachability could be carried via a combined MAC+IP route while a MAC only route may not be advertised at all. A single sequence number association with the MAC+IP route again implicitly assumes a fixed mapping between MAC and IP. A MAC move resulting in a new IP association for the host MAC breaks this assumption and results in a new MAC+IP route. If this new MAC+IP route independently assumes a new sequence number, this mobility attribute can no longer be used to determine most recent

host MAC reachability as opposed to the older existing MAC reachability.

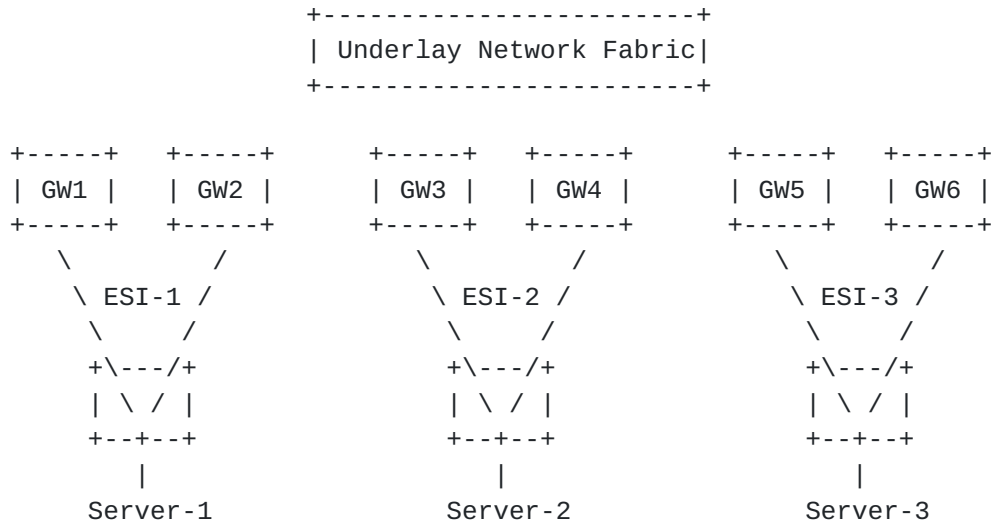


As an example, IP1-M1 is learnt locally at [GW1, GW2] and currently advertised to remote hosts with a sequence number N. Consider a scenario where a VM with MAC M1 is re-provisioned at server 2, however, as part of this re-provisioning, assigned a different IP address say IP7. [IP7, M1] is learnt as a new route at [GW3, GW4] and advertised to remote GWs with a sequence number of 0. As a result, L3 reachability to IP7 would be established across the overlay, however, MAC mobility procedure for MAC1 will not trigger as a result of this MAC-IP route advertisement. If an optional MAC only route is also advertised, sequence number associated with the MAC only route would trigger MAC mobility as per [RFC-7432](#). However, in the absence of an additional MAC only route advertisement, a single sequence number advertised with a combined MAC+IP route would not be sufficient to update MAC reachability across the overlay.

A MAC-IP sequence number assignment procedure needs to be defined to unambiguously determine the most recent MAC reachability in such a scenario without a MAC only route being advertised.

Further, GW1/GW2, on learning new reachability for [IP7, M1] via GW3/GW4 MUST probe and delete any local IPs associated with MAC M1, such as [IP1, M1] in the above example.

4. Multi-homed hosts via EVLAG



Consider an EVPN-IRB overlay network with hosts multi-homed to two or more leaf GW devices via an MC-LAG ESI. MAC and ARP entries learnt on a local ESI may also be synchronized across the multi-homing GW devices sharing this ESI. This MAC and ARP SYNC enables local switching of intra and inter subnet ECMP traffic flows from remote hosts. In other words, local MAC and ARP entries on a given Ethernet segment (ESI) may be learnt via local learning and / or sync from another GW device sharing the same ESI.

For a host that is multi-homed to multiple GW devices via an MC-LAG interface, local learning of host MAC and MAC-IP at each GW device is an independent asynchronous event, that is dependent on traffic flow and or ARP / ND response from the host hashing to a directly connected GW on the MC-LAG interface. As a result, sequence number mobility attribute value assigned to a locally learnt MAC or MAC-IP route (as per [RFC 7432](#)) at each device may not always be the same, depending on transient states on the device at the time of local learning.

As an example, consider a host VM that is deleted from ESI-2 and moved to ESI-1. It is possible for host to be learnt on say, GW1 following deletion of the remote route from [GW3, GW4], while being learnt on GW2 prior to deletion of remote route from [GW3, GW4]. If so, GW1 would process local host route learning as a new route and assign a sequence number of 0, while GW2 would process local host route learning as a remote to local move and assign a sequence number of N+1, N being the existing sequence number assigned at [GW3, GW4]. Inconsistent sequence numbers advertised from multi-homing devices

leads to various issues such as:

- o Ambiguity with respect to how the remote ToRs should handle paths with same ESI and different sequence numbers. A remote ToR may not program ECMP paths if it receives routes with different sequence numbers from a set of multi-homing GWs sharing the same ESI.
- o Breaks consistent route versioning across the network overlay that is needed for EVPN mobility procedures to work.

As an example, in this inconsistent state, GW2 would drop a remote route received for the same host with sequence number N (as its local sequence number is N+1), while GW1 would install it as the best route (as its local sequence number is 0).

There is need for a mechanism to ensure consistency of sequence numbers advertised from a set of multi-homing devices for EVPN mobility to work reliably.

In order to support mobility for multi-homed hosts using the sequence number mobility attribute, local MAC and MAC-IP routes MUST be advertised with the same sequence number by all GW devices that the ESI is multi-homed to. In other words, there is need for a mechanism to ensure consistency of sequence numbers advertised from a set of multi-homing devices for EVPN mobility to work reliably.

5. Design Considerations

To summarize, sequence number assignment scheme and implementation must take following considerations into account:

- o MAC+IP may be learnt on an ESI multi-homed to multiple GW devices, hence requires sequence numbers to be synchronized across multi-homing GW devices.
- o MAC only RT-2 is optional in an IRB scenario and may not necessarily be advertised in addition to MAC+IP RT-2
- o Single MAC may be associated with multiple IPs, i.e., multiple host IPs may share a common MAC
- o Host IP move could result in host moving to a new MAC, resulting in a new IP to MAC association and a new MAC+IP route.
- o Host MAC move to a new location could result in host MAC being associated with a different IP address, resulting in a new MAC to IP association and a new MAC+IP route

- o LOCAL MAC-IP learn is always accompanied by a LOCAL MAC learn, however, learning could happen in any order
- o Use cases discussed earlier that do not maintain a constant 1:1 MAC-IP mapping across moves could potentially be addressed by using separate sequence numbers associated with MAC and IP components of MAC+IP route. Maintaining two separate sequence numbers however adds significant overhead with respect to complexity, debugability, and backward compatibility. It is therefore goal of solution presented here to address these requirements via a single sequence number attribute.

6. Solution Components

This section goes over main components of the EVPN IRB mobility solution proposed in this draft. Later sections will go over exact sequence number assignment procedures resulting from concepts described in this section.

6.1 Sequence Number Inheritance

Main idea presented here is to view a LOCAL MAC-IP route as a child of the corresponding LOCAL MAC only route that inherits the sequence number attribute from the parent LOCAL MAC only route:

Mx-IPx -----> Mx (seq# = N)

As a result, both parent MAC and child MAC-IP routes share one common sequence number associated with the parent MAC route. Doing so ensures that a single sequence number attribute carried in a combined MAC+IP route represents sequence number for both a MAC only route as well as a MAC+IP route, and hence makes the MAC only route truly optional. As a result, optional MAC only route with its own sequence number is not required to establish most recent reachability for a MAC in the overlay network. Specifically, this enables a MAC to assume a different IP address on a move, and still be able to establish most recent reachability to the MAC across the overlay network via mobility attribute associated with the MAC+IP route advertisement. As an example, when Mx moves to a new location, it would result in LOCAL Mx being assigned a higher sequence number at its new location as per [RFC 7432](#). If this move results in Mx assuming a different IP address, IPz, LOCAL Mx+IPz route would inherit the new sequence number from Mx.

LOCAL MAC and LOCAL MAC-IP routes would typically be sourced from data plane learning and ARP learning respectively, and could get learnt in control plane in any order. Implementation could either

replicate inherited sequence number in each MAC-IP entry OR maintain a single attribute in the parent MAC by creating a forward reference LOCAL MAC object for cases where a LOCAL MAC-IP is learnt before the LOCAL MAC.

6.2 MAC Sharing

Further, for the shared MAC scenario, this would result in multiple LOCAL MAC-IP siblings inheriting sequence number attribute from a common parent MAC route:

```

Mx-IP1 -----
 |             |
Mx-IP2 -----
 .             |
 .             +----> Mx (seq# = N)
 .             |
Mx-IPw -----
 |             |
Mx-IPx -----

```

In such a case, a host-IP move to a different physical server would result in IP moving to a new MAC binding. A new MAC-IP route resulting from this move must now be advertised with a sequence number that is higher than the previous MAC-IP route for this IP, advertised from the prior location. As an example, consider a route Mx-IPx that is currently advertised with sequence number N from GW1. IPx moving to a new physical server behind GW2 results in IPx being associated with MAC Mz. A new local Mz-IPx route resulting from this move at GW2 must now be advertised with a sequence number higher than N. This is so that GW devices, including GW1, GW2, and other remote GW devices that are part of the overlay can clearly determine and program the most recent MAC binding and reachability for the IP. GW1 on receiving this new Mz-IPx route with sequence number say, N+1, for symmetric IRB case, would update IPx reachability via GW2 in forwarding, for asymmetric IRB case, would update IPx's ARP binding to Mz, clear and withdraw the stale Mx-IPx route with the lower sequence number.

This also implies that sequence number associated with local MAC Mz and all local MAC-IP children of Mz at GW2 must now be incremented to N+1, and re-advertised across the overlay. While this re-advertisement of all local MAC-IP children routes affected by the parent MAC route is an overhead, it avoids the need for two separate sequence number attributes to be maintained and advertised for IP and MAC components of MAC+IP RT-2. Implementation would need to be able to lookup MAC-IP routes for a given IP and update sequence number for

it's parent MAC and its MAC-IP children.

6.3 Multi-homing Mobility Synchronization

In order to support mobility for multi-homed hosts, local MAC and MAC-IP routes learnt on the shared ESI MUST be advertised with the same sequence number by all GW devices that the ESI is multi-homed to. This also applies to local MAC only routes. LOCAL MAC and MAC-IP may be learnt natively via data plane and ARP/ND respectively as well as via SYNC from another multi-homing GW to achieve local switching. Local and SYNC route learning can happen in any order. Local MAC-IP routes advertised by all multi-homing GW devices sharing the ESI must carry the same sequence number, independent of the order in which they are learnt. This implies:

- o On local or sync MAC-IP route learning, sequence number for the local MAC-IP route MUST be compared and updated to the higher value.
- o On local or sync MAC route learning, sequence number for the local MAC route MUST be compared and updated to the higher value.

If an update to local MAC-IP sequence number is required as a result of above comparison with sync MAC-IP route, it would essentially amount to a sequence number update on the parent local MAC, resulting in the inherited sequence number update on the MAC-IP route.

7. Requirements for Sequence Number Assignment

Following sections summarize sequence number assignment procedure needed on local and sync MAC and MAC-IP route learning events in order to accomplish the above.

7.1 LOCAL MAC-IP learning

A local Mx-IPx learning via ARP or ND should result in computation OR re-computation of parent MAC Mx's sequence number, following which the MAC-IP route Mx-IPx would simply inherit parent MAC's sequence number. Parent MAC Mx Sequence number should be computed as follows:

- o MUST be higher than any existing remote MAC route for Mx, as per [RFC 7432](#).
- o MUST be at least equal to corresponding SYNC MAC sequence number if one is present.
- o If the IP is also associated with a different remote MAC "Mz",

MUST be higher than "Mz" sequence number

Once new sequence number for MAC route Mx is computed as per above, all LOCAL MAC-IPs associated with MAC Mx MUST inherit the updated sequence number.

7.2 LOCAL MAC learning

Local MAC Mx Sequence number should be computed as follows:

- o MUST be higher than any existing remote MAC route for Mx, as per [RFC 7432](#).
- o MUST be at least equal to corresponding SYNC MAC sequence number if one is present.
- o Once new sequence number for MAC route Mx is computed as per above, all LOCAL MAC-IPs associated with MAC Mx MUST inherit the updated sequence number.

Note that the local MAC sequence number might already be present if there was a local MAC-IP learnt prior to the local MAC, in which case the above may not result in any change in local MAC's sequence number.

7.3 Remote MAC OR MAC-IP Update

On receiving a remote MAC OR MAC-IP route update associated with a MAC Mx with a sequence number that is higher than a LOCAL route for MAC Mx:

- o GW MUST trigger probe and deletion procedure for all LOCAL IPs associated with MAC Mx
- o GW MUST trigger deletion procedure for LOCAL MAC route for Mx

7.4 REMOTE (SYNC) MAC update

Corresponding local MAC Mx (if present) Sequence number should be re-computed as follows:

- o MUST be at least equal to corresponding SYNC MAC sequence number that is received.
- o If a LOCAL MAC sequence number is updated as a result of the above, all LOCAL MAC-IPs associated with MAC Mx MUST inherit the updated sequence number.

7.5 REMOTE (SYNC) MAC-IP update

If this is a SYNCed MAC-IP on a local ESI, it would also result in a derived SYNC MAC Mx route entry, as MAC only RT-2 advertisement is optional. Corresponding local MAC Mx (if present) Sequence number should be re-computed as follows:

- o MUST be at least equal to corresponding SYNC MAC sequence number that is received.
- o If a LOCAL MAC sequence number is updated as a result of the above, all LOCAL MAC-IPs associated with MAC Mx MUST inherit the updated sequence number.

7.6 Inter-op

In general, if all GW nodes in the overlay network follow the above sequence number assignment procedure, and the GW is advertising both MAC+IP and MAC routes, sequence number advertised with the MAC and MAC+IP routes with the same MAC would always be the same. However, an inter-op scenario with a different implementation could arise, where a GW implementation non-compliant with this proposal assigns and advertises independent sequence numbers to MAC and MAC+IP routes. To handle this case, if different sequence numbers are received for remote MAC+IP and corresponding remote MAC routes from a remote GW, sequence number associated with the remote MAC route should be computed as:

- o Highest of the all received sequence numbers with remote MAC+IP and MAC routes with the same MAC.
- o MAC sequence number would be re-computed on a MAC or MAC+IP route withdraw as per above.

A MAC and / or IP move to the local GW would now result in the MAC (and hence all MAC-IP) sequence numbers incremented from the above computed remote MAC sequence number.

8. Duplicate Host Detection

This section specifies additional considerations and requirements, incremental to the baseline duplicate MAC detection procedure described in [RFC 7432](#).

For all use cases where duplicate hosts have the same MAC, MAC is detected as duplicate via duplicate MAC detection procedure described in [RFC 7432](#). Corresponding MAC-IP routes with the same MAC do not require duplicate detection and MUST simply inherit the DUPLICATE

property from the corresponding MAC route. In other words, if a MAC route is in DUPLICATE state, all corresponding MAC-IP routes MUST also be treated as DUPLICATE. Duplicate detection procedure need only be applied to MAC routes.

However, due to misconfiguration, a situation may arise where hosts with different MACs are configured with the same IP. This scenario would not be detected by existing duplicate MAC detection procedure and would result in incorrect forwarding of routed traffic destined to this IP.

Such a situation, on LOCAL MAC-IP learning, would be detected as a move scenario via the following local MAC sequence number computation procedure described earlier in [section 5.1](#):

- o If the IP is also associated with a different remote MAC "Mz", MUST be higher than "Mz" sequence number

Such a move that results in sequence number increment on local MAC because of a remote MAC-IP route associated with a different MAC MUST be counted as an "IP move" against the "IP" independent of MAC. Duplicate detection procedure described in [RFC 7432](#) can now be applied to an "IP" entity independent of MAC. Once an IP is detected as DUPLICATE, corresponding MAC-IP route should be treated as DUPLICATE. Associated MAC routes and any other MAC-IP routes associated with this MAC should not be affected.

[8.1](#) Duplicate IP detection Procedure

- o If the IP is also associated with a different remote MAC "Mz", MUST be higher than "Mz" sequence number
- o On learning a LOCAL MAC-IP route Mx-IPx, check if there is an existing REMOTE route for IPx with a different MAC association, say, Mz-IPx
- o If so, count this as an "IP move" count for IPx, independent of the MAC
- o On learning a REMOTE MAC-IP route Mz-IPx, check if there is an existing LOCAL route for IPx with a different MAC association, say, Mx-IPx
- o If so, count this as an "IP move" count for IPx, independent of the MAC
- o Duplicate detection procedure described in [RFC 7432](#) can now be applied to IPx independent of MAC

- o If "N" IP moves are detected within "M" seconds for IPx, treat the corresponding LOCAL MAC-IP route for IPx as DUPLICATE
- o Apply duplicate handling on local MAC-IP route by alerting the operator and freezing EVPN RT-2 advertisements for it

A MAC-IP route SHOULD be treated as DUPLICATE if either of the following two conditions are met:

- o Corresponding MAC route is marked as DUPLICATE via existing duplicate detection procedure
- o Corresponding IP is marked as DUPLICATE via extended procedure described above

8.2 Duplicate Host Recovery

Once a MAC or IP is marked as DUPLICATE, corrective action must be taken to un-provision one of the duplicate MAC or IP. Once one of the duplicate hosts is un-provisioned, normal operation would not resume until the duplicate MAC ages out, following this correction, unless additional action is taken to speed up recovery.

This section lists possible additional corrective actions that could be taken to achieve fast recovery to normal operation.

8.2.1 Route Un-freezing CLI

Unfreezing the DUPLICATE MAC or IP via a CLI can be leveraged to automatically clear the duplicate MAC / IP entries following un-provisioning of the duplicate host:

- o If the duplicate host is un-provisioned at the location where it was NOT marked as DUPLICATE, unfreezing the route from the other location will result in automatic clearing of local routes on receiving higher sequence number routes following the un-freeze.
- o If the duplicate host is un-provisioned at the location where it was marked as DUPLICATE, unfreezing the route will trigger an advertisement with a higher sequence number to the other location. This would in-turn trigger re-learning of local route at the remote location, resulting in another advertisement with a higher sequence number from the remote location. Route at the local location would now be cleared on receiving this remote route advertisement and no re-learning would happen.

8.2.2 Route clearing CLI

Alternatively a CLI can be provided to clear the local route, to be executed AFTER the duplicate host is un-provisioned.

9. Security Considerations

10. IANA Considerations

11. References

11.1 Normative References

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", [RFC 7432](https://www.rfc-editor.org/rfc/7432), DOI 10.17487/RFC7432, February 2015, <<http://www.rfc-editor.org/info/rfc7432>>.

11.2 Informative References

12. Acknowledgements

Authors would like to thank Vibov Bhan and Patrice Brisset for feedback and comments through the process.

Authors' Addresses

Neeraj Malhotra
Cisco
EMail: nmalhotr@cisco.com

Ali Sajassi
Cisco
EMail: sajassi@cisco.com

Aparna Pattekar
Cisco
Email: apjoshi@cisco.com

Avinash Lingala
AT&T
Email: ar977m@att.com

Appendix A

An alternative approach considered was to associate two independent

sequence number attributes with MAC and IP components of a MAC-IP route. However, the approach of enabling IRB mobility procedures using a single sequence number associated with a MAC, as specified in this document was preferred for the following reasons:

- o Procedural overhead and complexity associated with maintaining two separate sequence numbers all the time, only to address scenarios with changing MAC-IP bindings is a big overhead for topologies where MAC-IP bindings never change.
- o Using a single sequence number associated with MAC is much simpler and adds no overhead for topologies where MAC-IP bindings never change.
- o Using a single sequence number associated with MAC is aligned with existing MAC mobility implementations. On other words, it is an easier implementation extension to existing MAC mobility procedure.

