Configuring VXLAN BGP-EVPN Null Route

This chapter contains the following sections:

About EVPN Null Route

A Distributed Denial of Service (DDoS) attack on a host in an EVPN Fabric consumes the network bandwidth resources and in turn impacts legitimate traffic to other hosts.

The DDoS attack can be from any of the following setups:

- Host connected to a leaf switch within the local site
- Host connected to a leaf switch in a remote site
- External networks such as WAN

The DDoS attack can be intra-subnets (MAC based) or inter-subnets (Host-based – IPv4/IPv6)

Null route filtering has been traditionally used in mitigating DDoS attacks especially in service provider networks.

A null route is a network route (routing table entry) that goes nowhere. Matching packets are dropped (ignored or redirected) rather than forwarded, acting as a kind of limited firewall. The act of using null routes is often called null route filtering.

NX-OS already has mechanisms to configure the null/drop route for IPv4/IPv6/MAC. The null route will be required to be configured on all VTEPs in the fabric.

For IPv4/IPv6 based attacks, use the following commands to configure an IPv4/IPv6 static route with null interface:

- **ip route x.x.x.x/y Null0**

- **ipv6 route X:X:X::X/Y Null0**

For MAC-based attacks, use the following command to configure MAC address with drop adjacency to drop the packets:

- **mac address-table static xxxx.yyyy.zzzz vlan <VLAN-ID> drop**

In a fabric with large number of VTEPs and across multiple sites, manually configuring and administering the drop route on all VTEPs is difficult task in the absence of Nexus Dashboard Fabric Controller (NDFC) or other Orchestrator.

The EVPN null routing feature is used when you do not have a way to configure and inject a null route from a central location such as with NDFC or other Orchestrators.

EVPN null routing feature enables a VTEP within the network to send Type-2 and Type-5 routes tagged with a specific community.

Other VTEPs (Borders and Leafs) in the single-site and multi-site can install an entry in MAC or IP (IPv4/IPv6) table such that any traffic destined to MAC or IP respectively is dropped at the Edge or leaf switch which prevents the usage of bandwidth within the site and across the site.

The programmed null route entry can be a Host IP (/32 or /128), a Prefix (VLSM) or a MAC.

Guidelines and Limitations for VXLAN BGP-EVPN Null Route

- A null route (static) MAC configuration must have matching static ARP/ND configuration which means you must not have a dynamic ARP/ND with MACs configured as null route MACs.
- If you use only L2-services (and has no configuration that can lead to dynamic ARP/ND learning) then a “mac drop” configuration alone is allowed. In all other cases, we require static ARP/ND configuration also along with the “mac drop” configuration.
- In case of vPC, the null route (MAC, mac-ip, prefix) must be configured on both vPC boxes (VMCT and PMCT). The behavior is undefined if this is not configured on both boxes. This rule applies for a host that is behind an orphan port on the vPC boxes. The configuration for null routing the orphan host must be made on both vPC boxes. The same holds good during unconfiguring the null route.
- The route-map must be applied on the remote VTEPs. This ingress Route-Map is important for Type-5 routes.
- No feature interaction with multicast traffic.
- When remote static is seen on a VTEP and if you want to configure the same MAC as a local static (static MAC with a valid interface or MAC set to drop/null route MAC), a syslog will be generated to warn about the duplicate configuration in the fabric that must be corrected. However, the configuration will not be rejected. The local static configuration holds precedence over a remote static configuration on that VTEP.
- If local static MAC with a valid interface is configured on a VTEP, and you want to convert this static MAC to a null route MAC on the same VTEP, the null route MAC takes effect.
- Though the remote dynamic MAC route permits any remote MAC route derived from MAC-IP route split to overwrite its entry, and propagate to MAC manager the remote static MAC route will no longer honor these derived MACs to overwrite its entry. As a result, the MAC entry remains unchanged until the remote static MAC is deleted.

- The null route MAC is another form of static MAC configuration only.

Configuring Static MAC

Before you begin

You can configure static drop MAC addresses. These static MAC addresses override dynamically learned MAC addresses on any interfaces.

SUMMARY STEPS

1. **configure terminal**
2. **mac address-table static** *mac-address* **vlan** *vlan-id* {[**drop**| **interface**{*type slot/port*} | **port-channel** *number*]}
3. **exit**

DETAILED STEPS

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**mac address-table static** *mac-address* **vlan** *vlan-id* {[**drop**| **interface**{*type slot/port*}	**port-channel** *number*]}
		Example:
		`switch(config)# mac address-table static 3001.3010.99aa vlan 3001 drop`
		`switch(config)#`
			Specifies a static MAC address to add to the Layer 2 MAC address table.

Step 3	**exit**
	Example:
	`switch# exit`
	`switch#`
		Exits the configuration mode.

Configuring ARP/ND

You can configure ARP/ND host on IPv4/IPv6 route for the corresponding SVI.

Before you begin

Ensure to configure static MAC-IP configuration on the switch where MAC is configured as drop entry. This will avoid MAC-IP mobility and ensures both DROP MAC and MAC-IP are originated from same VTEP.

SUMMARY STEPS

1. **configure terminal**
2. **interface** *vlan-number*
3. **vrf member** *vrf-name*
4. **no ip redirects**
5. **ip address** *address*
6. **ipv6 address** *address*
7. **ipv6 neighbor address** *ipv6address mac_addr*
8. **no ipv6 redirects**
9. **ip arp address** *ipaddr mac_addr*
10. **fabric forwarding mode anycast-gateway**

DETAILED STEPS

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**interface** *vlan-number*
	Example:
	`switch(config)# interface Vlan 3001`
	`switch(config-if)#`
		Specifies the VLAN interface.

Step 3	**vrf member** *vrf-name*
	Example:
	`switch(config-if)# vrf member cgw_3001_3050`
	`switch(config-if)#`
		Assigns the VLAN interface to the tenant VRF.

Step 4	**no ip redirects**
	Example:
	`switch(config-if)# no ip redirects`
	`switch(config-if)#`
		Disables the IPv4 redirects.

Step 5	**ip address** *address*
	Example:
	`switch(config-if)# ip address 30.1.0.1/16`
	`switch(config-if)#`
		Specifies the IP address.

Step 6	**ipv6 address** *address*
	Example:
	`switch(config-if)# ipv6 address 2001:3001::1/64`
	`switch(config-if)#`
		Specifies the IPv6 address.

Step 7	**ipv6 neighbor address** *ipv6address mac_addr* **Example:**	Configures static IPv6 neighbor.

	Command or Action	Purpose
	`switch(config-if)# ipv6 neighbor 2001:3001::99 3001.3010.99aa` `switch(config-if)#`	

Step 8	**no ipv6 redirects**
	Example:
	`switch(config-if)# no ipv6 redirects`
	`switch(config-if)#`
		Disables the IPv6 redirects.

Step 9	**ip arp address** *ipaddr mac_addr*
	Example:
	`switch(config-if)# ip arp 30.1.0.99 3001.3010.99aa`
	`switch(config-if)#`
		Associates an IP address with a MAC address as a static entry.

Step 10	**fabric forwarding mode anycast-gateway**
	Example:
	`switch# fabric forwarding mode anycast-gateway`
	`switch#`
		Associates SVI with anycast gateway under VLAN configuration mode.

Configuring Prefix-Null Route on Local VTEP

On a local VTEP where the Null route is configured, configure route-map to set blackhole community on static route and redistribute into BGP.

SUMMARY STEPS

1. **configure terminal**
2. **vrf context** *vrf-name*
3. **ip route** {<*ip*>/*mask*} **Null0 tag** <*tag-number*> or **ip route** {<*ipv6*>/*mask*} **Null0 tag** <*tag-number*>
4. **route-map** *map-name* [**permit** | **deny**] [*seq*]
5. **match tag** <*tag-number*>
6. **set weight** *value*
7. **set community blackhole**
8. **router bgp** *as-number*
9. **vrf** *vrf-name*
10. **address-family ipv4/ipv6 unicast**
11. **redistribute static route-map** *route-map name*

DETAILED STEPS

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

	Command or Action	Purpose

Step 2	**vrf context** *vrf-name*
	Example:
	`switch(config)# vrf context tenant-0001`
	`switch(config-vrf)#`
		Configures the tenant VRF.

Step 3	**ip route** {<*ip*>/*mask*} **Null0 tag** <*tag-number*> or **ip route** {<*ipv6*>/*mask*} **Null0 tag** <*tag-number*>
	Example:
	For IPv4
	`switch(config-vrf)# ip route 50.1.0.0/24 Null0 tag 6666`
	`switch(config-vrf)#`
	For IPv6
	`switch(config-vrf)# ipv6 route 50::1:0/120 Null0 tag 6666`
	`switch(config-vrf)#`
		Configures static-route for destination prefix with Null0 nexthop and matching tag.

Step 4	**route-map** *map-name* [**permit** | **deny**] [*seq*]
	Example:
	`switch(config)# route-map SET_BHC permit 10`
	`switch(config-route-map)#`
		Creates a route map or enters route-map configuration mode for an existing route map. Use seq to order the entries in a route map.

Step 5	**match tag** <*tag-number*>
	Example:
	`switch(config-route-map)# match tag 6666`
	`switch(config-route-map)#`
		Matches the routes with the configured tag.

Step 6	**set weight** *value*
	Example:
	`switch (config-route-map)# set weight 65535`
	`switch(config-route-map)#`
		Sets the weight for the incoming route with blackhole community. we recommend to set the **set weight** value to maximum value, to give the highest precedence to the null routes. The maximum value of **set weight** is 65535.

Step 7	**set community blackhole**
	Example:
	`switch(config-route-map)# set community blackhole`
	`switch(config-route-map)#`
		Sets the community as Blackhole (well-known community).

Step 8	**router bgp** *as-number*
	Example:
	`switch(config)# router bgp 100`
	`switch(config-router)#`
		Enables a routing process. The range of as-num is 1–65535.

Step 9	**vrf** *vrf-name*
	Example:
	`switch(config-router)# vrf tenant-0001`
	`switch(config-router-vrf)#`
		Configures the tenant VRF.

	Command or Action	Purpose

Step 10	**address-family ipv4/ipv6 unicast**
	Example:
	`switch(config-router-vrf)# address-family ipv4 unicast`
	`switch(config-router-vrf-af)#`
		Configure the IPv4/IPv6 address family. This configuration is required for IPv4/IPv6 over VXLAN with IPv4/IPv6 underlay.

Step 11	**redistribute static route-map** *route-map name*
	Example:
	`switch(config-router-vrf-af)# redistribute static route-map SET_BHC`
	`switch(config-router-vrf-af)#`
		Redistributes the prefix-null static route into BGP using the configured route-map.

Configuring RPM Route-Map on Remote VTEP

Before you begin

On remote VTEP, use a community-list and route-map to give precedence to the null routes:

SUMMARY STEPS

1. **configure terminal**
2. **ip community-list standard** <*community-list-name*> **seq** <*seq-number*> **permit blackhole**
3. **route-map** *map-name***[permit | deny]** <*seq-number*>
4. **match community** <*community-list*>
5. **set weight** *value*
6. **route-map** *map-name***permit** <*seq-number*>
7. **router bgp** *as-number*
8. **route-map** *route-map* {**in** | **out**}

DETAILED STEPS

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`switch# configure terminal`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip community-list standard** <*community-list-name*> **seq** <*seq-number*> **permit blackhole**
	Example:
	`switch (config)# ip community-list standard BH seq 10 permit blackhole`
	`switch(config)#`
		Configures a community list and permits routes that have the well-known "blackhole" community value.
		Beginning with Cisco NX-OS Release 10.3(2)F, the blackhole (well-known community) is added to the existing IP community list.

	Command or Action	Purpose

Step 3	**route-map** *map-name***[permit | deny]** <*seq-number*>
	Example:
	`switch(config)# route-map PREFER_BHC permit 10`
	`switch(config-route-map)#`
		Enters route-map configuration mode

Step 4	**match community** <*community-list*>
	Example:
	`switch(config-route-map)# match community BH`
	`switch(config-route-map)#`
		The BGP routes are matched using the community list.

Step 5	**set weight** *value*
	Example:
	`switch (config-route-map)# set weight 65535`
	`switch(config-route-map)#`
		Sets the weight for the incoming route with blackhole community. we recommend to set the **set weight** value to maximum value, to give the highest precedence to the null routes. The maximum value of **set weight** is 65535.

Step 6	**route-map** *map-name***permit** <*seq-number*>
	Example:
	`switch(config-route-map)# route-map PREFER_BHC permit 20`
	`switch(config-route-map)#`
		Configures a route-map with a fallback permit clause to allow other routes.

Step 7	**router bgp** *as-number*
	Example:
	`switch(config)# router bgp 100`
	`switch(config-router)#`
		Enables a routing process. The range of as-num is from 1 to 65535.

Step 8	**route-map** *route-map* {**in** | **out**} **Example:** `switch(config-router-neighbor-af)# route-map PREFER_BHC in`	Applies the route map to the neighbor in the configured direction.

Configuration Example for Null Route

The following example shows how to set the local/remote configuration on prefix-null and MAC/MAC-IP drop routes:

Configuration – Prefix Null

On local VTEP (Border leaf switch) where the Type-5 null route is to be advertised, perform the following steps:

1. Configure static IPv4/IPv6 address with Null0 adjacency

```
vrf context tenant-0001
  vni 3100001
  ip route 50.1.0.0/24 Null0 tag 6666
  ipv6 route 50::1:0/120 Null0 tag 6666
```

2. Configure route-map to set null route community on static route and redistribute into BGP

```
route-map SET_BHC permit 10
  match tag 6666
  set community blackhole
router bgp 100
  router-id 10.1.0.21
  vrf tenant-0001
    address-family ipv4 unicast
      redistribute static route-map SET_BHC
    address-family ipv6 unicast
      redistribute static route-map SET_BHC
```

On all other remote VTEPs, perform the following steps:

1. Configure route-map to match the null route community and set weight to highest value to ensure null route is always preferred.

```
ip community-list standard BH seq 10 permit blackhole
route-map PREFER_BHC permit 10
  match community BH
  set weight 65535
route-map PREFER_BHC permit 20
router bgp 100
 router-id 10.1.0.13
 address-family l2vpn evpn
 template peer LEAF_to_FABRIC_IBGP_OVERLAY
    remote-as 100
    address-family l2vpn evpn
      send-community
      send-community extended
      route-map PREFER_BHC in
```

Configuration – MAC/MAC-IP Drop

On local VTEP where Type-2 null route is to be advertised, perform the following steps:

Note In vPC or VMCT setup, the same local VTEP configuration must be applied on both peer switches if advertising blackholes route from the vPC switch. There is no consistency checker for misconfiguration on vPC peer.

1. Configure static MAC address with drop adjacency

```
mac address-table static 0013.e001.0001 vlan 2 drop
```

2. Configure static ARP/ND neighbor for same address

```
interface Vlan2
  no shutdown
  vrf member tenant-0001
  ip address 5.0.63.254/18
  ipv6 address 5::3f7f/114
  ipv6 neighbor 5::17fe 0013.e001.0001
  no ipv6 redirects
  ip arp 5.0.23.254 0013.e001.0001
  fabric forwarding mode anycast-gateway
```

On all other remote VTEPs, perform the following step:

1. Configure route-map to match the blackhole community and set weight to highest value to ensure null route is always preferred.

```
ip community-list standard BH seq 10 permit blackhole
route-map PREFER_BHC permit 10
  match community BH
  set weight 65535
route-map PREFER_BHC permit 20
router bgp 100
 router-id 10.1.0.13
 address-family l2vpn evpn
 template peer LEAF_to_FABRIC_IBGP_OVERLAY
    remote-as 100
    address-family l2vpn evpn
      send-community
      send-community extended
      route-map PREFER_BHC in
  neighbor 10.1.0.31
    inherit peer LEAF_to_FABRIC_IBGP_OVERLAY
```

Verifying EVPN Null Route Configuration

To display the EVPN null route configuration information, enter one of the following commands:

Command	Purpose
show bgp l2vpn evpn	Displays routing table information.
show ip arp static vlan <*vlan-id*> **vrf** <*vrf-name*>	Displays local ARP information.
show ip arp static remote vlan <*vlan-id*> **vrf** <*vrf-name*>	Displays remote ARP information.
show ip adjacency vlan <*vlan-id*> **detail vrf** <*vrf-name*>	Displays local adjacency information.
show ipv6 icmp neighbour static remote [*vlan* <*id*>] [*vrf* <*name*>]	Displays remote static neighbor information.
show mac address-table static vlan <*vlan-id*>	Displays local/remote MAC information.
show ip community-list *name*	Displays information about a IP community list.
show route-map *name*	Displays information about a route map.

The following example shows Type-2 EVPN Route sample output for the **show bgp l2vpn evpn** command:

```
switch# show bgp l2vpn evpn 1111.1111.1111
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 53.53.53.53:32769    (L2VNI 1000002)
BGP routing table entry for [2]:[0]:[0]:[48]:[1111.1111.1111]:[32]:[100.100.100.51]/272,
version 23
Paths: (1 available, best #1)
Flags: (0x000102) (high32 00000000) on xmit-list, is not in l2rib/evpn
Multipath: eBGP iBGP
  Advertised path-id 1
  Path type: local, path is valid, is best path, no labeled nexthop, has esi_gw
  AS-Path: NONE, path locally originated
```

```
53.53.53.53 (metric 0) from 0.0.0.0 (53.53.53.53)
Origin IGP, MED not set, localpref 100, weight 32768
Received label 1000002 1000100
Community: Blackhole
Extcommunity: RT:23456:1000002 RT:23456:1000100 ENCAP:8
Router MAC:0476.b0f0.8157
Path-id 1 advertised to peers:
111.111.54.1
```

The following example shows Type-5 EVPN Route (sent) sample output for the **show bgp l2vpn evpn** command:

```
switch# sh bgp ipv4 uni 44.44.44.0 vrf 100
BGP routing table information for VRF 100, address family IPv4 Unicast
BGP routing table entry for 44.44.44.0/24, version 6
Paths: (1 available, best #1)
Flags: (0x80c0002) (high32 0x000020) on xmit-list, is not in urib, exported, has label
vpn: version 5, (0x00000000100002) on xmit-list
local label: 492287

Advertised path-id 1, VPN AF advertised path-id 1
Path type: redist, path is valid, is best path, no labeled nexthop, is extd
AS-Path: NONE, path locally originated
0.0.0.0 (metric 0) from 0.0.0.0 (44.44.44.44)
Origin incomplete, MED 0, localpref 100, weight 32768
Community: blackhole
Extcommunity: RT:23456:1000100

VRF advertise information:
Path-id 1 not advertised to any peer

VPN AF advertise information:
Path-id 1 not advertised to any peer

switch# sh bgp l2 e 44.44.44.0
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 53.53.53.53:4 (L3VNI 1000100)
BGP routing table entry for [5]:[0]:[0]:[24]:[44.44.44.0]/224, version 5
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn
Multipath: eBGP iBGP

Advertised path-id 1
Path type: local, path is valid, is best path, no labeled nexthop, has esi_gw
Gateway IP: 0.0.0.0
AS-Path: NONE, path locally originated
53.53.53.53 (metric 0) from 0.0.0.0 (53.53.53.53)
Origin incomplete, MED 0, localpref 100, weight 32768
Received label 1000100
Community: blackhole
Extcommunity: RT:23456:1000100 ENCAP:8 Router MAC:0476.b0f0.8157

Path-id 1 advertised to peers:
111.111.54.1
```

The following example shows Type-5 EVPN Route (received) sample output for the **show bgp l2vpn evpn** command:

```
switch# sh bgp l2 e 44.44.44.0
BGP routing table information for VRF default, address family L2VPN EVPN
Route Distinguisher: 53.53.53.53:4
BGP routing table entry for [5]:[0]:[0]:[24]:[44.44.44.0]/224, version 2
Paths: (1 available, best #1)
Flags: (0x000002) (high32 00000000) on xmit-list, is not in l2rib/evpn, is not in HW
```

```
Multipath: eBGP iBGP

Advertised path-id 1
Path type: external, path is valid, is best path, no labeled nexthop, has esi_gw
Imported to 2 destination(s)
Imported paths list: 100 L3-1000100
Gateway IP: 0.0.0.0
AS-Path: 4241653625 , path sourced external to AS
53.53.53.53 (metric 2) from 111.111.53.1 (53.53.53.53)
Origin incomplete, MED 0, localpref 100, weight 0
Received label 1000100
Community: blackhole
Extcommunity: RT:11000:1000100 Route-Import:53.53.53.53:100
Source AS:4241653625:0 SOO:50529024:00000000 ENCAP:8
Router MAC:0476.b0f0.8157
Path-id 1 not advertised to any peer

switch# show bgp ipv4 uni 44.44.44.0 vrf 100
BGP routing table information for VRF 100, address family IPv4 Unicast
BGP routing table entry for 44.44.44.0/24, version 3
Paths: (1 available, best #1)
Flags: (0x8008001a) (high32 00000000) on xmit-list, is in urib, is best urib route, is in
HW
vpn: version 3, (0x00000000100002) on xmit-list

Advertised path-id 1, VPN AF advertised path-id 1
Path type: external, path is valid, is best path, no labeled nexthop, in rib, has esi_gw

Imported from 53.53.53.53:4:[5]:[0]:[0]:[24]:[44.44.44.0]/224
AS-Path: 4241653625 , path sourced external to AS
53.53.53.53 (metric 2) from 111.111.53.1 (53.53.53.53)
Origin incomplete, MED 0, localpref 100, weight 0
Received label 1000100
Community: blackhole
Extcommunity: RT:11000:1000100 Route-Import:53.53.53.53:100
Source AS:4241653625:0 SOO:50529024:00000000 ENCAP:8
Router MAC:0476.b0f0.8157

VRF advertise information:
Path-id 1 not advertised to any peer
```