

DRNI+EVPN组网及配置介绍

H3C交换机技术专题篇

DRNI+EVPN

EVPN大家还记得吗？不记得的小伙伴快点击上方话题复习一下！

EVPN是一种基于Overlay技术的二层网络互联技术，可以将不同的数据中心站点互联起来，降低数据中心的成本、灵活扩充数据中心业务。目前，EVPN不仅广泛应用于数据中心网络，在园区接入网络、广域网、运营商网络中也具有一定的应用。

DRNI技术可以实现设备级的冗余和流量负载分担，我们可以将EVPN和DRNI结合部署，把EVPN组网中的两台Leaf、Border设备连接起来虚拟成一台设备，使用该虚拟设备作为Leaf、Border，可以避免Leaf、Border单点故障对网络造成影响，从而提高EVPN网络的可靠性。

那DRNI和EVPN可以结合在一起吗？当然可以啦，让我们一起来看一下DRNI+EVPN组网应该如何规划配置吧~



DRNI+EVPN

部署方案说明

在介绍具体配置之前，我们先通过一个EVPN Leaf、Spine、Border三级组网大概说明一下，DRNI+EVPN方案具体是如何实现的。



如图，Leaf作为VTEP或者EVPN网关，实现服务器之间二三层流量的转发，通过部署DRNI使服务器双归接入VTEP或EVPN网关，避免单条链路故障导致虚拟机无法访问网络；

Spine作为路由反射器，负责在Leaf、Border之间反射路由；

Border作为边界网关，实现数据中心与外部网络的互通；ED作为本数据中心的边缘设备，通过VXLAN DCI隧道实现本数据中心与其他数据中心的互通；通过部署DRNI，避免设备故障导致无法访问外部网络或其他数据中心。

在DRNI+EVPN组网中，流量是如何转发的呢？

服务器通过DR口接入Leaf设备，在Leaf设备上查表，通过VXLAN隧道完成服务器之间二三层流量的转发，当服务器需要访问外部网络时，Leaf通过VXLAN隧道将流量转发到Border，再由Border查找路由表完成服务器对外部网络的访问；

同一DRNI系统的两台Leaf/Border具有相同的虚拟地址（通过evpn drni group命令配置），对外表现为一台虚拟设备。其他Leaf/Border使用该地址与这台虚拟设备自动建立VXLAN隧道；

同一DRNI系统的两台Leaf/Border使用不同的地址作为BGP对等体邻居，分别与作为反射器的Spine建立BGP EVPN邻居。利用Underlay网络的等价路由机制，其他Leaf/Border可以将发往虚拟Leaf/Border地址的流量同时发送到两台Leaf/Border，从而实现负载分担和冗余备份。

DRNI+EVPN

具体配置介绍

前面我们介绍了整个DRNI+EVPN组网的部署方案思路，接下来我们通过一个简化组网的HCL模拟器实验，详细介绍一下DRNI+EVPN组网的具体配置。

如图，Switch A、Switch B、Switch D是与服务器连接的VTEP设备（即Leaf）。Switch A和Switch B通过DRNI虚拟为一台VTEP设备，Switch A和Switch B之间通过IPL同步MAC地址和ARP信息，以确保两台VTEP上的MAC地址和ARP信息保持一致。Switch C（即Spine）作为路由反射器在Switch A、Switch B、Switch D之间反射路由。



Switch A和Switch B均通过以太网链路与下行的Access设备相连，Access设备上分别下挂虚拟机VM1、VM2，要求在连接每一台Access设备的链路间跨设备建立二层聚合接口，避免以太网链路故障导致虚拟机无法访问网络。虚拟机VM1、VM3同属于VXLAN 10，通过EVPN实现不同站点间的二层互通；VM2属于VXLAN 20，与VM1、VM2之间三层互通。

接下来我们来一起进行配置吧~

首先在交换机上配置路由协议，使得各交换机的接口IP地址（包括Loopback接口IP地址）之间Underlay路由可达。本次实验我们以OSPF路由协议为例。

```
[配置Switch A/B/C/D分别在OSPF分布接口所在网段的路由]
[SwitchA] ospf 1 router-id 1.1.1.1
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 1.1.1.1 0.0.0.0 --Loopback口地址
[SwitchA-ospf-1-area-0.0.0.0] network 1.2.3.4 0.0.0.0 --虚拟VTEP地址
[SwitchA-ospf-1-area-0.0.0.0] network 11.1.1.0 0.0.0.255 --互联接口地址

[SwitchB] ospf 1 router-id 2.2.2.2
[SwitchB-ospf-1] area 1
[SwitchB-ospf-1-area-0.0.0.0] network 1.2.3.4 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 2.2.3.4 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 12.1.1.0 0.0.0.255

[SwitchC] ospf 1 router-id 3.3.3.3
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 3.3.3.3 0.0.0.0
[SwitchC-ospf-1-area-0.0.0.0] network 11.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 12.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255

[SwitchD] ospf 1 router-id 4.4.4.4
[SwitchD-ospf-1] area 0
[SwitchD-ospf-1-area-0.0.0.0] network 4.4.4.4 0.0.0.0
[SwitchD-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
```

接下来我们在Switch A、Switch B上开启EVPN支持DRNI功能，使两台设备虚拟为一台VTEP设备。

```
[Switch A]
开启EVPN的分布式聚合模式，并配置虚拟VTEP地址为1.2.3.4。
[SwitchA] evpn drni group 1.2.3.4
配置DR系统。
[SwitchA] drni system-mac 0001-0002-0003
[SwitchA] drni system-number 1
[SwitchA] drni system-priority 10
[SwitchA] drni restore-delay 180
[SwitchA] drni keepalive ip destination 60.1.1.2 source 60.1.1.1
创建二层聚合接口3，配置为IRPP1，并关闭报文入口与静态MAC地址表项匹配检查功能。
[SwitchA] interface bridge-aggregation 3
[SwitchA-Bridge-Aggregation3] link-aggregation mode dynamic
[SwitchA] interface gigabitEthernet 1/0/3
[SwitchA-Ten-GigabitEthernet1/0/3] port link-aggregation group 3
[SwitchA] interface bridge-aggregation 3
[SwitchA-Bridge-Aggregation3] port drni intra-portal-port 1
[SwitchA-Bridge-Aggregation3] undo mac-address static source-check enable -- 这条命令根据组网不需要配置，但实际组网中必须配置
创建二层聚合接口4和5，配置为DR口
[SwitchA] interface bridge-aggregation 4
[SwitchA-Bridge-Aggregation4] link-aggregation mode dynamic
[SwitchA-Bridge-Aggregation4] quit
[SwitchA] interface gigabitEthernet 1/0/1
[SwitchA-Ten-GigabitEthernet1/0/1] port link-aggregation group 4
[SwitchA] interface bridge-aggregation 4
[SwitchA-Bridge-Aggregation4] port drni group 4
[SwitchA] interface bridge-aggregation 5
[SwitchA-Bridge-Aggregation5] link-aggregation mode dynamic
[SwitchA] interface gigabitEthernet 1/0/2
[SwitchA-Ten-GigabitEthernet1/0/2] port link-aggregation group 5
[SwitchA-Bridge-Aggregation5] port drni group 5
将所有参与EVPN业务的接口配置为保留接口（G1/0/4口是Keepalive1）
[SwitchA] drni mad exclude interface loopback 0
[SwitchA] drni mad exclude interface gigabitEthernet 1/0/4
[SwitchA] drni mad exclude interface vlan-interface 11
[SwitchA] drni mad exclude interface vsi-interface 1
[SwitchA] drni mad exclude interface vsi-interface 2
[SwitchA] drni mad exclude interface vsi-interface 3
```

Switch B的配置和上面Switch A的配置类似，需要注意的是，drni system-number要和Switch A配置得不一样，但evpn drni group地址两台设备必须配置成相同的，以达到虚拟成一台VTEP设备的目的。

DRNI系统建立好后，我们可以开始进行Overlay的配置。在Switch A、Switch B和Switch D上配置EVPN，使VTEP之间/关联VXLAN隧道、通告MAC/IP的可达性等。

```
[Switch A]
[SwitchA] bgp 200
[SwitchA-bgp-default] peer 3.3.3.3 as-number 200
[SwitchA-bgp-default] peer 3.3.3.3 connect-interface loopback 0
[SwitchA-bgp-default-evpn] address-family l2vpn evpn
[SwitchA-bgp-default-evpn] peer 3.3.3.3 enable

[Switch B]
[SwitchB] bgp 200
[SwitchB-bgp-default] peer 3.3.3.3 as-number 200
[SwitchB-bgp-default] peer 3.3.3.3 connect-interface loopback 0
[SwitchB-bgp-default-evpn] address-family l2vpn evpn
[SwitchB-bgp-default-evpn] peer 3.3.3.3 enable

[Switch D]
[SwitchD] bgp 200
[SwitchD-bgp-default] peer 3.3.3.3 as-number 200
[SwitchD-bgp-default] peer 3.3.3.3 connect-interface loopback 0
[SwitchD-bgp-default-evpn] address-family l2vpn evpn
[SwitchD-bgp-default-evpn] peer 3.3.3.3 enable
```

此外还需要配置Spine Switch C作为路由反射器在Switch A、Switch B、Switch D之间反射路由。

```
[Switch C]
[SwitchC-bgp-default] group evpn
[SwitchC-bgp-default] peer 1.1.1.1 group evpn
[SwitchC-bgp-default] peer 2.2.2.2 group evpn
[SwitchC-bgp-default] peer 4.4.4.4 group evpn
[SwitchC-bgp-default] peer evpn as-number 200
[SwitchC-bgp-default] peer evpn connect-interface loopback 0
[SwitchC-bgp-default-evpn] address-family l2vpn evpn
[SwitchC-bgp-default-evpn] undo policy vpn-target
[SwitchC-bgp-default-evpn] peer evpn reflect-client --反射器
```

BGP EVPN配置好后，下一步需要在VTEP上创建VSI，配置EVPN实例等，这里需要注意，两台DRNI设备Switch A和Switch B的EVPN配置必须一致。

```
[Switch A]
开启L2VPN能力
关闭远端MAC地址和远端ARP自动学习功能
[SwitchA] vxlan tunnel mac-learning disable
[SwitchA] vxlan tunnel arp-learning disable
配置EVPN的全局MAC地址为0002-0003-0004
[SwitchA] evpn global-mac 2-3-4
在VSI实例vpna下创建EVPN实例
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] evpn encapsulation vxlan
配置自动生成EVPN实例的RD和RT
[SwitchA-vsi-vpna-evpn-vxlan] route-distinguisher auto
[SwitchA-vsi-vpna-evpn-vxlan] vpn-target auto
创建VXLAN 10
在VSI实例vpnb下创建EVPN实例
[SwitchA] vsi vpb
[SwitchA-vsi-vpb] evpn encapsulation vxlan
[SwitchA-vsi-vpb-evpn-vxlan] route-distinguisher auto
[SwitchA-vsi-vpb-evpn-vxlan] vpn-target auto
创建VXLAN 20
[SwitchA-vsi-vpb] vxlan 20
[SwitchB]
[SwitchB] l2vpn enable
[SwitchB] vxlan tunnel mac-learning disable
[SwitchB] vxlan tunnel arp-learning disable
[SwitchB] evpn global-mac 2-3-4
[SwitchB] vsi vpna
[SwitchB-vsi-vpna] evpn encapsulation vxlan
[SwitchB-vsi-vpna-evpn-vxlan] route-distinguisher auto
[SwitchB-vsi-vpna-evpn-vxlan] vpn-target auto
[SwitchB] vsi vpb
[SwitchB-vsi-vpb] evpn encapsulation vxlan
[SwitchB-vsi-vpb-evpn-vxlan] route-distinguisher auto
[SwitchB-vsi-vpb-evpn-vxlan] vpn-target auto
[SwitchB-vsi-vpb] vxlan 20
[SwitchD]
[SwitchD] l2vpn enable
[SwitchD] vxlan tunnel mac-learning disable
[SwitchD] vxlan tunnel arp-learning disable
[SwitchD] vsi vpna
[SwitchD-vsi-vpna] evpn encapsulation vxlan
[SwitchD-vsi-vpna-evpn-vxlan] route-distinguisher auto
[SwitchD-vsi-vpna-evpn-vxlan] vpn-target auto
[SwitchD-vsi-vpna] vxlan 10
```

VTEP设备上的配置是类似的，Spine上只做转发，无需配置实例。

VM1和VM2、VM3的网关都在Leaf上，我们需要在Switch A/B/D上配置分布式网关，这里Switch A和Switch B的配置也是完全一致的。

```
[Switch A&B]
配置L3VPN的RD和RT。
[SwitchA] ip vpn-instance vpna
[SwitchA-vpn-instance-vpna] route-distinguisher 1:1
[SwitchA-vpn-instance-vpna] address-family ipv4
[SwitchA-vpn-ipv4-vpna] vpn-target 2:2
[SwitchA-vpn-ipv4-vpna] quit
[SwitchA-vpn-instance-vpna] address-family evpn
[SwitchA-vpn-evpn-vpna] vpn-target 1:1
配置VSI虚接口VSI-interface1。
[SwitchA] interface vsi-interface 1
[SwitchA-Vsi-interface1] ip binding vpn-instance vpna
[SwitchA-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchA-Vsi-interface1] mac-address 1-1-1
[SwitchA-Vsi-interface1] distributed-gateway local
[SwitchA-Vsi-interface1] local-proxy-arp enable
配置VSI虚接口VSI-interface2。
[SwitchA] interface vsi-interface 2
[SwitchA-Vsi-interface2] ip binding vpn-instance vpna
[SwitchA-Vsi-interface2] ip address 10.1.1.1 255.255.255.0
[SwitchA-Vsi-interface2] distributed-gateway local
[SwitchA-Vsi-interface2] local-proxy-arp enable
创建VSI虚接口VSI-interface3，在该接口上配置VPN实例vpna对应的L3VNI为1000。
[SwitchA] interface vsi-interface 3
[SwitchA-Vsi-interface3] ip binding vpn-instance vpna
[SwitchA-Vsi-interface3] l3-vni 1000
配置VXLAN 10所在的VSI实例和接口VSI-interface1关联。
[SwitchA] vsi vpna
[SwitchA-vsi-vpna] gateway vsi-interface 1
配置VXLAN 20所在的VSI实例和接口VSI-interface2关联。
[SwitchA] vsi vpb
[SwitchA-vsi-vpb] gateway vsi-interface 2
[SwitchD]
配置L3VPN的RD和RT。
[SwitchD] ip vpn-instance vpna
[SwitchD-vpn-instance-vpna] route-distinguisher 1:1
[SwitchD-vpn-instance-vpna] address-family ipv4
[SwitchD-vpn-ipv4-vpna] vpn-target 2:2
[SwitchD-vpn-instance-vpna] address-family evpn
[SwitchD-vpn-evpn-vpna] vpn-target 1:1
配置VSI虚接口VSI-interface1。
[SwitchD] interface vsi-interface 1
[SwitchD-Vsi-interface1] ip binding vpn-instance vpna
[SwitchD-Vsi-interface1] ip address 10.1.1.1 255.255.255.0
[SwitchD-Vsi-interface1] mac-address 1-1-1
[SwitchD-Vsi-interface1] distributed-gateway local
[SwitchD-Vsi-interface1] local-proxy-arp enable
创建VSI虚接口VSI-interface3，在该接口上配置VPN实例vpna对应的L3VNI为1000。
[SwitchD] interface vsi-interface 3
[SwitchD-Vsi-interface3] ip binding vpn-instance vpna
[SwitchD-Vsi-interface3] l3-vni 1000
配置VXLAN 10所在的VSI实例和接口VSI-interface1关联。
[SwitchD] vsi vpna
[SwitchD-vsi-vpna] gateway vsi-interface 1
```

然后在Switch A、Switch B和Switch D的下行端口（即AC口）上配置以太网服务实例和相应的匹配规则，用来识别用户网络中的报文所属的VXLAN。

```
[Switch A]
接入服务器的接口Bridge-Aggregation4上创建以太网服务实例1000，该实例用来匹配VLAN 2的数据帧。
[SwitchA] interface bridge-aggregation 4
[SwitchA-Bridge-Aggregation4] service-instance 1000
[SwitchA-Bridge-Aggregation4-srv1000] encapsulation s-vid 2
配置以太网服务实例1000与VSI实例vpna关联。
[SwitchA-Bridge-Aggregation4-srv1000] xconnect vsi vpna
接入服务器的接口Bridge-Aggregation5上创建以太网服务实例1000，该实例用来匹配VLAN 3的数据帧。
[SwitchA] interface bridge-aggregation 5
[SwitchA-Bridge-Aggregation5] service-instance 1000
[SwitchA-Bridge-Aggregation5-srv1000] encapsulation s-vid 3
配置以太网服务实例1000与VSI实例vpnb关联。
[SwitchA-Bridge-Aggregation5-srv1000] xconnect vsi vpb
[Switch B]
接入服务器的接口Bridge-Aggregation4上创建以太网服务实例1000，该实例用来匹配VLAN 2的数据帧。
[SwitchB] interface bridge-aggregation 4
[SwitchB-Bridge-Aggregation4] service-instance 1000
[SwitchB-Bridge-Aggregation4-srv1000] encapsulation s-vid 2
配置以太网服务实例1000与VSI实例vpna关联。
[SwitchB-Bridge-Aggregation4-srv1000] xconnect vsi vpna
接入服务器的接口Bridge-Aggregation5上创建以太网服务实例1000，该实例用来匹配VLAN 3的数据帧。
[SwitchB] interface bridge-aggregation 5
[SwitchB-Bridge-Aggregation5] service-instance 1000
[SwitchB-Bridge-Aggregation5-srv1000] encapsulation s-vid 3
配置以太网服务实例1000与VSI实例vpnb关联。
[SwitchB-Bridge-Aggregation5-srv1000] xconnect vsi vpb
[Switch D]
接入服务器的接口Ten-GigabitEthernet1/0/1上创建以太网服务实例1000，该实例用来匹配VLAN 2的数据帧。
[SwitchD] interface ten-gigabitEthernet 1/0/1
[SwitchD-Ten-GigabitEthernet1/0/1] service-instance 1000
[SwitchD-Ten-GigabitEthernet1/0/1-srv1000] encapsulation s-vid 2
配置以太网服务实例1000与VSI实例vpna关联。
[SwitchD-Ten-GigabitEthernet1/0/1-srv1000] xconnect vsi vpna
```

到这里一个简单的DRNI+EVPN组网就搭建好啦，我们先来看一下VTEP之间隧道建立的情况：

```
[SwitchA]display interface Tunnel
Tunnel0
Current state: UP
Line protocol state: UP
Description: Tunnel0 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 1.2.3.4, destination 4.4.4.4 ---本端group ip与远端地址
Tunnel protocol/transport UDP_VXLAN/IP

[SwitchB]display interface Tunnel
Tunnel0
Current state: UP
Line protocol state: UP
Description: Tunnel0 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 4.4.4.4, destination 1.2.3.4 ---本端地址与远端group ip
Tunnel protocol/transport UDP_VXLAN/IP

[SwitchD]display interface Tunnel
Tunnel0
Current state: UP
Line protocol state: UP
Description: Tunnel0 Interface
Bandwidth: 64 kbps
Maximum transmission unit: 1464
Internet protocol processing: Disabled
Last clearing of counters: Never
Tunnel source 4.4.4.4, destination 1.2.3.4 ---本端地址与远端group ip
Tunnel protocol/transport UDP_VXLAN/IP
```

Switch A和Switch B上都是通过evpn drni group命令配置的虚地址和远端Leaf建立的隧道，远端Leaf Switch D上建立的Tunnel目的地址也是虚地址1.2.3.4，我们在Switch D上看它收到的路由，可以看到Switch D从Switch A/B上收到的2类和5类路由下一跳都是group ip，对Switch D而言，Switch A/B已经虚拟成一台设备。

```
[SwitchD]display bgp l2vpn evpn
BGP local router ID is 4.4.4.4
Status codes: * - valid, > - best, d - dampened, h - history
s - suppressed, S - stale, i - internal, e - external
a - additional-path
Origin: i - IGP, e - EGP, ? - incomplete
Route distinguisher: 1:1(vpna)
Total number of routes: 5
* > i Network: [2][0][48][28ca-d77c-0702][32][10.1.1.10]/136
NextHop: 1.2.3.4 LocPrf: 100
PrefVal: 0 OutLabel: NULL
MED: 0
Path/Ogn: i
* > i Network: [2][0][48][28ca-e269-0b02][32][10.1.2.20]/136
NextHop: 1.2.3.4 LocPrf: 100
PrefVal: 0 OutLabel: NULL
MED: 0
Path/Ogn: i
* > Network: [5][0][24][10.1.1.0]/80 ---本地
NextHop: 0.0.0.0 LocPrf: 100
PrefVal: 32768 OutLabel: NULL
MED: 0
Path/Ogn: i
* i Network: [5][0][24][10.1.1.0]/80
NextHop: 1.2.3.4 LocPrf: 100
PrefVal: 0 OutLabel: NULL
MED: 0
Path/Ogn: i
* > Network: [5][0][24][10.1.2.0]/80
NextHop: 1.2.3.4 LocPrf: 100
PrefVal: 0 OutLabel: NULL
MED: 0
Path/Ogn: i
```

除此以外，我们在Switch A/B上查看VSI信息，可以看到设备会自动在IPL上创建AC，并将其与VSI关联。这个Type为DRNI的AC，就是设备自动创建的。

```
[SwitchA]display l2vpn vsi verbose
VSI Name: Auto_L3VNI1000_3
...
Gateway Interface : VSI-interface 3
VXLAN ID : 1000

VSI Name: vpna
...
Gateway Interface : VSI-interface 1
VXLAN ID : 10

Tunnels:
Tunnel Name Link ID State Type Flood proxy
Tunnel0 0x5000000 UP Auto Disabled
ACs:
AC Link ID State Type
BAGG4 srv1000 0 Up Manual
BAGG3 srv2 1 Up Dynamic (DRNI)

VSI Name: vpb
...
Gateway Interface : VSI-interface 2
VXLAN ID : 20
ACs:
AC Link ID State Type
BAGG5 srv1000 0 Up Manual
BAGG3 srv3 1 Up Dynamic (DRNI)
```

以上就是一个简单的DRNI+EVPN组网最基本的配置啦，看完后大家是不是还有一些疑问呢，比如Type为DRNI的AC是什么？Border应该怎么配置？如果终端只能单挂接入应该怎么办？别着急，下期我们就来一一解答，不要错过哦~

老规矩，点击文末“阅读原文”即可获取本期金手指-HCL模拟器工程文件，在下期上难度之前，再好好复习一下吧！

神秘代码: drni

要是有问题or建议记得留言

扫码关注关注我们哦

—end—

