



# ECMP Support in OpenStack Contrail Environments

Avi Technical Reference (v20.1)

Copyright © 2020

# ECMP Support in OpenStack Contrail Environments

[view online](#)

Avi Vantage can manage load balancing capacity for a virtual service by dynamically scaling it out or in on additional or fewer Service Engines (SEs). By default, the primary SE for the virtual service co-ordinates distribution of traffic flow amongst the secondary SEs, including itself. On OpenStack with Contrail, Avi Vantage can take advantage of Contrail's ECMP support and manage the orchestration of ECMP routes as part of virtual service placement.

ECMP functionality can be homed at the following locations:

1. The upstream edge router (e.g., Juniper MX or such)
2. The Contrail vRouter on the host hypervisor.

## Example

In Neutron, the VIP `19.1.1.7` is associated with the interface port of the SE named `Avi-se-czpey`.

```
{%cli%} root@dc11-cfg-1:~# neutron port-list +-----+-----+-----+-----+
+-----+-----+-----+-----+ | id | name | mac_address | fixed_ips |
+-----+-----+-----+-----+
+-----+-----+-----+-----+ | dc4695c6-6d16-4ee0-ab05-774537a33ab2 | Avi-
Data...| 02:dc:46:95:c6:6d | {"subnet_id": "a834986a-385e-4616-9d9c-6b91cfaa51e0", "ip_address": "19.1.1.6"} |||
{"subnet_id": "a834986a-385e-4616-9d9c-6b91cfaa51e0", "ip_address": "19.1.1.7"} | ... +-----+
+-----+-----+-----+-----+ + {%endcli%}
```

In Contrail, the VIP `19.1.1.7` has the same linkage to port `dc4695c6-6d16-4ee0-ab05-774537a33ab2` and one route to the SE VM running on hypervisor `172.16.11.101`.

## Virtual Service Scaleout

Using the Avi CLI, setting `scaleout_ecmp` to `True` enables a scaleout of `vs2` as follows:

```
{%cli%} demo:11-1-1-7: > configure virtualservice vs2 demo:11-1-1-7: virtualservice> scaleout_ecmp demo:11-1-1-7:
virtualservice> save +-----+-----+-----+-----+ | Field | Value |
+-----+-----+-----+-----+ | uuid | virtualservice-2f8770d0-20cc-477f-
9419-b206b9624389 || name | vs2 || scaleout_ecmp | True || vip[1] || vip_id | 0 || ip_address | 19.1.1.7 || enabled | True ||
network_ref | 02b846e6-cd68-4c84-b7ab-3709946ed28b || port_uuid | c0e4cb54-8706-4bed-97bb-271a603f858f ||
subnet_uuid | a834986a-385e-4616-9d9c-6b91cfaa51e0 || subnet | 19.1.1.0/24 || auto_allocate_ip | True ||
auto_allocate_floating_ip | False | ... +-----+
+-----+-----+-----+-----+ + {%endcli%}
```

After the scaleout, `vs2` runs on two SEs, `Avi-se-czpey` and `Avi-se-pnjbb`, as shown in the below summary output.

```
{%cli%} demo:11-1-1-7: > show virtualservice vs2 summary +-----+-----+-----+-----+ | Field |
Value | +-----+-----+-----+-----+ | oper_status || state | OPER_UP || percent_ses_up | 100
|| vip_summary[1] || vip_id | 0 || oper_status || state | OPER_UP || service_engine[1] || ref | Avi-se-czpey || primary | True ||
standby | False || service_engine[2] || ref | Avi-se-pnjbb || primary | False || standby | False || num_se_requested | 2 ||
num_se_assigned | 2 | +-----+-----+-----+-----+ + {%endcli%}
```

In Neutron, the VIP `19.1.1.7` is now associated with the interface ports of SE `Avi-se-czpey` and `Avi-se-pnjbb` as well.

```
{%cli%} root@dc11-cfg-1:~# neutron port-list +-----+-----+-----+
+-----+-----+-----+ | id | name | mac_address | fixed_ips |
+-----+-----+-----+
+-----+-----+-----+ | dc4695c6-6d16-4ee0-ab05-774537a33ab2 | Avi-
Data...| 02:dc:46:95:c6:6d | {"subnet_id": "a834986a-385e-4616-9d9c-6b91cfaa51e0", "ip_address": "19.1.1.6"} |||||
{"subnet_id": "a834986a-385e-4616-9d9c-6b91cfaa51e0", "ip_address": "19.1.1.7"} |
| 492ee72a-c647-490f-9648-77124e1cf0b6 | Avi-Data...| 02:49:2e:e7:2a:c6 | {"subnet_id": "a834986a-385e-4616-9d9c-
6b91cfaa51e0", "ip_address": "19.1.1.4"} ||||| {"subnet_id": "a834986a-385e-4616-9d9c-6b91cfaa51e0", "ip_address": "19.1.1.7"} |
+-----+-----+-----+
+-----+-----+-----+ {%endcli%}
```

In Contrail, the VIP 19.1.1.7 has the same linkage to ports dc4695c6-6d16-4ee0-ab05-774537a33ab2 and 492ee72a-c647-490f-9648-77124e1cf0b6, and two routes to the SE VM running on hypervisors 172.16.11.101 and 172.16.11.102.

If there are multiple SE instances for a VIP address, then the vRouter will receive multiple XMPP routes towards the final destination. The XMPP routes will have different route distinguishers (RDs) to keep them distinct, and they have different next-hops and MPLS labels to identify the different SE instances.

## Flow Resiliency During Scale Out/In

The flow is 5-tuple: src-IP, src-port, dst-IP, dst-port, and protocol. Routers do a hash of the 5-tuple to pick a path to use. Contrail vRouter uses flow-tables to preserve the chosen SE destination for a flow. Thus, even when the number of SE instances in the ECMP group changes, a flow will never move.

## Floating-IP ECMP

Using Avi CLI, associate a floating IP address with virtual service vs2 as follows:

```
{%cli%} demo:11-1-1-7: > configure virtualservice vs2 demo:11-1-1-7: virtualservice> vip vip_id 0 demo:11-1-1-7: virtualservice:
vip> auto_allocate_floating_ip Overwriting the previously entered value for auto_allocate_floating_ip demo:11-1-1-7:
virtualservice:vip> floating_subnet_uuid 010a4d99-794c-42a9-948c-080014463217 demo:11-1-1-7: virtualservice:vip> save
demo:11-1-1-7: virtualservice> save +-----+-----+-----+ | Field |
Value | +-----+-----+-----+ | uuid | virtualservice-2f8770d0-
20cc-477f-9419-b206b9624389 || name | vs2 || scaleout_ecmp | True || vip[1] || vip_id | 0 || ip_address | 19.1.1.7 || enabled |
True || network_ref | 02b846e6-cd68-4c84-b7ab-3709946ed28b || port_uuid | c0e4cb54-8706-4bed-97bb-271a603f858f |
| subnet_uuid | a834986a-385e-4616-9d9c-6b91cfaa51e0 || subnet | 19.1.1.0/24 || auto_allocate_ip | True ||
auto_allocate_floating_ip | False || floating_ip | 10.1.11.4 || auto_allocate_floating_ip | True || floating_subnet_uuid | 010a4d99-
794c-42a9-948c-080014463217 | ... +-----+-----+-----+ {%
endcli%}
```

In Contrail, the `floating_ip` is linked to the interface ports of both SEs and two routes to the SE VM running on hypervisors 172.16.11.101 and 172.16.11.102.

## Virtual Service Traffic Stats

```
{%cli%} demo:11-1-1-7: > show virtualservice vs2 detail filter disable_aggregate se | grep connections_handled |
connections_handled | 102 || connections_handled | 49 | {%endcli%}
```

Search Routes

Routing Instance: default-domainsandbox1:public-user1-net:public-user1-net

Prefix: Prefix

Limit: 50 Routes

Search Reset

Prefix	Protocol	Source	Next hop	Label	Security Group	Origin VN
Routing Table: default-domainsandbox1:public-user1-net:public-user1-net:evpn.0 (2 Routes)						
2-00-0-02:49:2e:e7:2a:c6:10.1.11.4	XMPP	dc11-compute-2	172.16.11.102	28	-	default-domainsandbox1:public-user1-net
2-00-0-02:dc:46:95:c6:6d:10.1.11.4	XMPP	dc11-compute-1	172.16.11.101	34	-	default-domainsandbox1:public-user1-net

Figure A. ECMP FIP linkage.

Search Routes

Routing Instance: default-domainsandbox1:public-user1-net:public-user1-net

Prefix: Prefix

Limit: 50 Routes

Search Reset

Prefix	Protocol	Source	Next hop	Label	Security Group	Origin VN
Routing Table: default-domainsandbox1:public-user1-net:public-user1-net:evpn.0 (2 Routes)						
2-00-0-02:49:2e:e7:2a:c6:10.1.11.4	XMPP	dc11-compute-2	172.16.11.102	28	-	default-domainsandbox1:public-user1-net
2-00-0-02:dc:46:95:c6:6d:10.1.11.4	XMPP	dc11-compute-1	172.16.11.101	34	-	default-domainsandbox1:public-user1-net
Routing Table: default-domainsandbox1:public-user1-net:public-user1-net:inet.0 (2 Routes)						
10.1.11.4/32	XMPP	dc11-compute-1	172.16.11.101	33	-	default-domainsandbox1:public-user1-net
10.1.11.4/32	XMPP	dc11-compute-2	172.16.11.102	27	-	default-domainsandbox1:public-user1-net

Figure B. ECMP FIP routing table.

```

10.10.10.100:9100/instance-ip/38338969-0680-4085-82a0-c9e253c3a23e
{
  "group": "cloud-admin-group",
  "group_access": 7
},
- fq_name: [
  "38338969-0680-4085-82a0-c9e253c3a23e"
],
instance_ip_local_ip: false,
instance_ip_mode: "active-active",
instance_ip_secondary: false,
name: "38338969-0680-4085-82a0-c9e253c3a23e",
display_name: "38338969-0680-4085-82a0-c9e253c3a23e",
uuid: "38338969-0680-4085-82a0-c9e253c3a23e",
instance_ip_address: "19.1.1.7",
- virtual_machine_interface_refs: [
- {
- to: [
- "default-domain",
- "demo",
- "Avi-Data#cluster-f7950d50-7bdc-40d7-bd93-70a608a28aa0#cloud-c254bb54-af69-415e-9e07-4b28871cafcd#4b9f33a6-72f8-11e7-b5b8-0204342b4b81"
],
href: "http://10.10.10.100:9100/virtual-machine-interface/dc4695c6-6d16-4ee0-ab05-774537a33ab2",
attr: null,
uuid: "dc4695c6-6d16-4ee0-ab05-774537a33ab2"
},
- {
- to: [
- "default-domain",
- "demo",
- "Avi-Data#cluster-f7950d50-7bdc-40d7-bd93-70a608a28aa0#cloud-c254bb54-af69-415e-9e07-4b28871cafcd#eb4c6d2c-72fa-11e7-b5b8-0204342b4b81"
],
href: "http://10.10.10.100:9100/virtual-machine-interface/492ee72a-c647-490f-9648-77124e1cf0b6",
attr: null,
uuid: "492ee72a-c647-490f-9648-77124e1cf0b6"
}
]
}
    
```

Figure C. ECMP linkage.

```

10.10.10.100:9100/instance-ip/38338969-0680-4085-82a0-c9e253c3a23e

instance_ip_local_ip: false,
instance_ip_secondary: false,
name: "38338969-0680-4085-82a0-c9e253c3a23e",
display_name: "38338969-0680-4085-82a0-c9e253c3a23e",
uuid: "38338969-0680-4085-82a0-c9e253c3a23e",
instance_ip_address: "19.1.1.7",
- virtual_machine_interface_refs: [
  - {
    - to: [
      "default-domain",
      "demo",
      "Avi-Data#cluster-f7950d50-7bdc-40d7-bd93-70a608a28aa0#cloud-c254bb54-af69-415e-9e07-4b28871cafcd#4b9f33a6-72f8-11e7-b5b8-0204342b4b81"
    ],
    href: "http://10.10.10.100:9100/virtual-machine-interface/dc4695c6-6d16-4ee0-ab05-774537a33ab2",
    attr: null,
    uuid: "dc4695c6-6d16-4ee0-ab05-774537a33ab2"
  },
],

```

Figure D. Non-ECMP linkage

Monitor > Infrastructure > Control Nodes > dc11-cfg-1

Details Peers Routes Console Alarms

Search Routes

Click here to expand

Routes

Prefix	Protocol	Source	Next hop	Label	Security Group	Origin VN
Routing Table: bgp.evpn.0 (1 Route)						
2-172.16.11.101:4-0-02:dc4695c6:6d16:19.1.1.7	XMPP	dc11-compute-1	172.16.11.101	34	-	default-domain:demo:frontend

Figure E. Non-ECMP route 1.

Monitor > Infrastructure > Control Nodes > dc11-cfg-1

Details Peers Routes Console Alarms

Search Routes

Click here to expand

Routes

Prefix	Protocol	Source	Next hop	Label	Security Group	Origin VN
Routing Table: bgp.evpn.0 (2 Routes)						
2-172.16.11.101:4-0-02:dc4695c6:6d16:19.1.1.7	XMPP	dc11-compute-1	172.16.11.101	34	-	default-domain:demo:frontend
2-172.16.11.102:1-0-02:49:2e:e7:2a:c6:19.1.1.7	XMPP	dc11-compute-2	172.16.11.102	28	-	default-domain:demo:frontend

Figure F. ECMP route 2.