ECMP and Asymmetric Return Path Case Study

This article demonstrates asymmetric routing: return path on a different interface.

1 What is ECMP?

Equal Cost Multi-Path (ECMP) is a mechanism that allows multiple routes to the same destination with different next-hops in the routing. ECMP also load-balances routed traffic over those multiple next-hops. (*Load-balancing is not within the scope of this document.)*

The FortiGate ECMP algorithm is IP source hash based on the IP address before network address translation (NAT) takes place. ECMP is supported for OSPF and static routing. ECMP only works for routes that are sourced by the same routing protocol (i.e. Static routes or OSPF).

Configuring ECMP using the following CLI command:

```
config system settings
  set ecmp-max-paths  (10 is default)
end
```

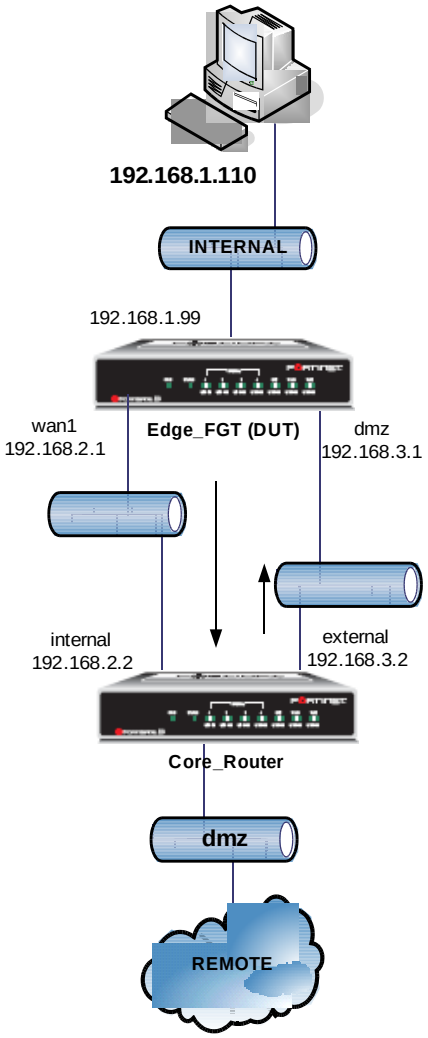
2 Network configuration

The diagram on the next page shows an example ECMP network configuration.

In this configuration the Core_Router is another FortiGate unit configured with a specific static route to 192.168.1.110/32 out on its 192.168.2.2 interface and asymmetric routing enabled to allow return traffic to egress the 192.168.3.2 interface.

```
Core_Router # get router info routing-table all

S*      0.0.0.0/0 [10/0] via 192.168.183.254, dmz
S       192.168.1.0/24 [10/0]   via 192.168.2.1, internal
S       192.168.1.110/32 [10/0] via 192.168.3.1, external
C       192.168.2.0/24 is directly connected, internal
C       192.168.3.0/24 is directly connected, external
C       192.168.182.0/23 is directly connected, dmz
```



3 Device under test

```
Edge_FGT # get system status
Version: Fortigate-60 3.00,build0730
Hostname: Edge_FGT
Operation Mode: NAT
Current virtual domain: root
Branch point: 730
MR/Patch Information: MR7 Patch 1
```

4 Configuration excerpt (all other is default) of Edge_FGT (DUT)

```
config router static
    edit 1
        set device "wan1"
        set gateway 192.168.2.2
    next
    edit 2
        set device "dmz"
        set gateway 192.168.3.2
        set priority 100
    next
end
```

The above settings force traffic to egress on wan1, and accept packets on dmz (rpf).

```
config firewall policy
    edit 1
        set srcintf "internal"
        set dstintf "wan1"
            set srcaddr "all"
            set dstaddr "all"
        set action accept
        set schedule "always"
            set service "ANY"
    next
end
```

```
Edge_FGT # get system setting
opmode              : nat
ecmp-max-paths      : 10
```

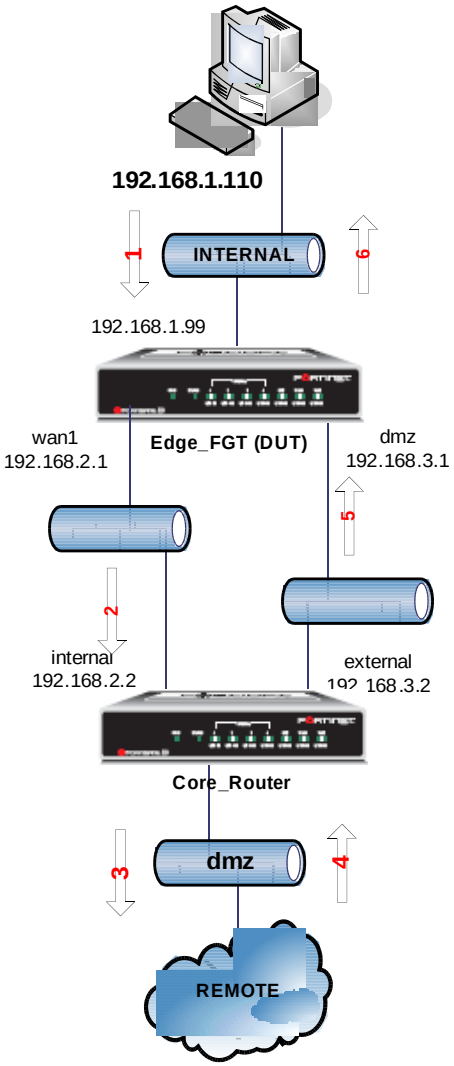
```
Edge_FGT # get router info routing-table all
S*      0.0.0.0/0 [10/0] via 192.168.2.2, wan1
                         [10/0] via 192.168.3.2, dmz
C       192.168.1.0/24 is directly connected, internal
C       192.168.2.0/24 is directly connected, wan1
C       192.168.3.0/24 is directly connected, dmz
```

5 Start the test: Ping 192.168.4.1 from the PC (source 192.168.1.110)

Sniffer trace on both devices and packet numbers:

```
Edge_FGT # diagnose sniffer packet  any "icmp" 4
1.315069 internal in 192.168.1.110 -> 192.168.4.1: icmp: echo request  □ 1
1.315141 wan1 out 192.168.1.110 -> 192.168.4.1: icmp: echo request □ 2
1.382564 dmz in 192.168.4.1 -> 192.168.1.110: icmp: echo reply □ 5

Core_Router # diagnose sniffer  packet  any "icmp" 4
1.254148 internal in 192.168.1.110 -> 192.168.4.1: icmp: echo request □ 2
1.254199 dmz out 192.168.182.102 -> 192.168.4.1: icmp: echo request □ 3
1.321123 dmz in 192.168.4.1 -> 192.168.182.102: icmp: echo reply □ 4
1.321181 external out 192.168.4.1 -> 192.168.1.110: icmp: echo reply □ 5
```



6 Debug flow and session information

```
Edge_FGT # diagnose debug flow filter addr 192.168.1.110
Edge_FGT # diagnose debug flow show console  enable
Edge_FGT # diagnose debug flow trace start 1000
Edge_FGT # diag debug enable

Edge_FGT # id=20085 trace_id=335 msg="vd-root received a packet(proto=1, 192.168.1.110:512-
>192.168.4.1:8) from internal."
id=20085 trace_id=335 msg="allocate a new session-00000142"
id=20085 trace_id=335 msg="find a route: gw-192.168.2.2 via wan1"
id=20085 trace_id=335 msg="Allowed by Policy-1:"
id=20085 trace_id=336 msg="vd-root received a packet(proto=1, 192.168.4.1:512->192.168.1.110:0)
from dmz."
id=20085 trace_id=336 msg="Find an existing session, id-00000142, reply direction"
id=20085 trace_id=336 msg="find a route: gw-192.168.1.110 via internal"
diagnose sys session clearid=20085 trace_id=337 msg="vd-root received a
packet(proto=1,192.168.1.110:512->192.168.4.1:8) from internal."
id=20085 trace_id=337 msg="Find an existing session, id-00000142, original direction"
id=20085 trace_id=337 msg="enter fast path"
id=20085 trace_id=338 msg="vd-root received a packet(proto=1, 192.168.4.1:512->192.168.1.110:0)
from dmz."
id=20085 trace_id=338 msg="Find an existing session, id-00000142, reply direction"
id=20085 trace_id=338 msg="enter fast path"



Edge_FGT # diag sys  session filter src 192.168.1.110
Edge_FGT # diag sys session list

session info: proto=1 proto_state=00 expire=59 timeout=3600 flags=00000000 sockf
lag=00000000 sockport=0 av_idx=0 use=3
bandwidth=0/sec guaranteed_bandwidth=0/sec      traffic=0/sec   prio=0  ha_id=0 hakey=879
tunnel=/state=may_dirty
statistic(bytes/packets/err): org=17160/286/0 reply=17160/286/0 tuples=2
orgin->sink: org pre->post, reply pre->post dev=5->3/3->5 gwy=192.168.2.2/192.168.1.110
hook=pre dir=org act=noop 192.168.1.110:512->192.168.4.1:8(0.0.0.0:0)
hook=post dir=reply act=noop 192.168.4.1:512->192.168.1.110:0(0.0.0.0:0)
misc=20002 policy_id=1 auth_info=0 ids=0xc5bfc000 vd=0 serial=00000142 tos=ff/ff app=0
total session 1
```