IMPROVING IPV6 SUPPORT ON FREEBSD

Luiz Amaral

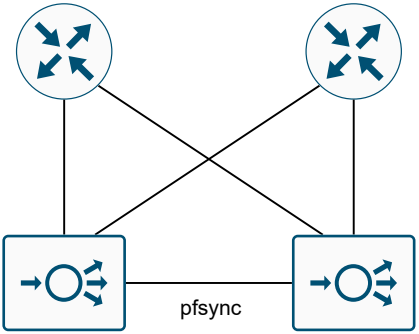
HOW DID WE GET HERE?

HOW DID WE GET HERE?

- FreeBSD is our OS of choice for load balancers
- We are trying to reduce the internal IPv4 usage as much as possible

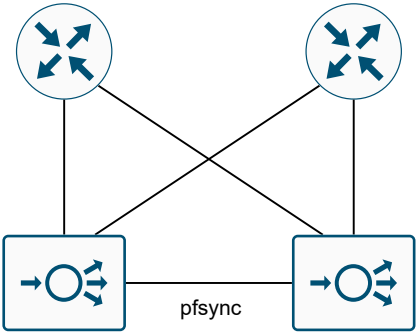
HOW DID WE GET HERE?

- FreeBSD is our OS of choice for load balancers
- We are trying to reduce the internal IPv4 usage as much as possible



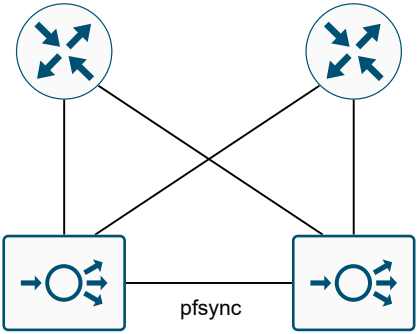
HOW DID WE GET HERE?

- FreeBSD is our OS of choice for load balancers
- We are trying to reduce the internal IPv4 usage as much as possible
- Good candidates were the dedicated pfsync interfaces and uplink interfaces towards core routers



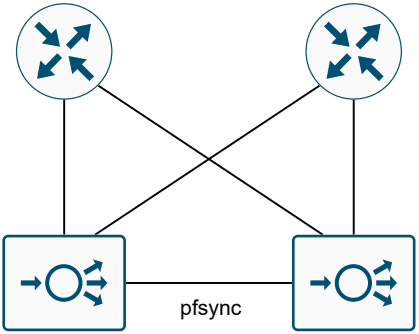
HOW DID WE GET HERE?

- FreeBSD is our OS of choice for load balancers
- We are trying to reduce the internal IPv4 usage as much as possible
- Good candidates were the dedicated pfsync interfaces and uplink interfaces towards core routers
- Remove IPv4 addresses from these interfaces



HOW DID WE GET HERE?

- FreeBSD is our OS of choice for load balancers
- We are trying to reduce the internal IPv4 usage as much as possible
- Good candidates were the dedicated pfsync interfaces and uplink interfaces towards core routers
- Remove IPv4 addresses from these interfaces
- Profit?



WELL, NOT SO FAST

WELL, NOT SO FAST

pfsync had no support for transport over IPv6 at all:

```
$ sudo ifconfig pfsync0 syncpeer fe80::1
ifconfig: error in parsing address string: Address family for
    hostname not supported
```

WELL, NOT SO FAST

pfsync had no support for transport over IPv6 at all:

```
$ sudo ifconfig pfsync0 syncpeer fe80::1
ifconfig: error in parsing address string: Address family for
    hostname not supported
```

IPV6 SUPPORT FOR PFSYNC

With a little bit of open-source magic:

```
813c5b75e680 - pfsync: prepare code to accommodate AF_INET6
   family
485be9798a75 - pfsync: replace struct pfsync_pkt with int flags
6fc7fc2dbb2b - pfsync: transport over IPv6
```

- https://reviews.freebsd.org/D36277
- https://reviews.freebsd.org/D36294
- https://reviews.freebsd.org/D40102
  With contributions from Naman Sood (FreeBSD Foundation)

IPV6 SUPPORT FOR PFSYNC

- Starting with FreeBSD 14.0 (being released around 2023-11-14)
- Unicast and multicast support, just like the IPv4 implementation
- Link-local addresses are sufficient. No need for ULA or GUA addresses

IPV6 SUPPORT FOR PFSYNC

- Starting with FreeBSD 14.0 (being released around 2023-11-14)
- Unicast and multicast support, just like the IPv4 implementation
- Link-local addresses are sufficient. No need for ULA or GUA addresses

```
$ sudo ifconfig pfsync0 syncdev vtnet0
$ sudo ifconfig pfsync0 syncpeer ff12::f0
$ ifconfig pfsync0
pfsync0: flags=1000041<UP,RUNNING,LOWER_UP> metric 0 mtu 1500
        options=0
        syncdev: vtnet0 syncpeer: ff12::f0%vtnet0 maxupd: 128
           defer: off version: 1400
        syncok: 1
        groups: pfsync
```

HOW ABOUT IPV4 OVER IPV6 NEXTHOP?

On Kernel:

```
$ route add 172.16.0.0/18 -inet6 2a12:6e40:1:1::1
$ netstat -nr
...
Destination        Gateway            Flags      Netif Expire
172.16.0.0/18      2a12:6e40:1:1::1   UGS       vtnet1
```

On BIRD:

```
bird> show route
Table master4:
172.18.0.0/18         unicast [static1 14:05:36.220] ! (200)
        via 2a12:6e40:1:1::1 on vtnet1
```

OH BOY, HERE WE GO AGAIN

Checking the logs, we get a hint of what is going on:

```
bird[2430]: KRT: Error sending route 172.18.0.0/18 to kernel:
    Network is unreachable
bird[2430]: KRT: Invalid route received (172.16.0.0/18) -
    missing gateway
```

OH BOY, HERE WE GO AGAIN

Checking the logs, we get a hint of what is going on:

```
bird[2430]: KRT: Error sending route 172.18.0.0/18 to kernel:
    Network is unreachable
bird[2430]: KRT: Invalid route received (172.16.0.0/18) -
    missing gateway
```

We tried Netlink a few days later, but it was also failing to import routes into BIRD.

BIRD BUGFIXES

```
d61505b0 - BSD: IPv4 over IPv6 nexthop support on FreeBSD
--- a/sysdep/bsd/krt-sock.c
+++ b/sysdep/bsd/krt-sock.c
@@ -314,7 +314,7 @@ krt_send_route(struct krt_proto *p, int cmd, rte *
    e)
        if (ipa_is_link_local(gw))
         _I0(gw) = 0xfe800000 | (i->index & 0x0000ffff);

-       sockaddr_fill(&gate, af, gw, NULL, 0);
+       sockaddr_fill(&gate, (ipa_is_ip4(gw) ? AF_INET : AF_INET6), gw,
     NULL, 0);
@@ -469,7 +469,7 @@ krt_read_route(struct ks_msg *msg, struct
    krt_proto *p, int scan)
     idst  = ipa_from_sa(&dst);
    imask = ipa_from_sa(&mask);
-   igate = (gate.sa.sa_family == dst.sa.sa_family) ? ipa_from_sa(&gate
    ) : IPA_NONE;
+   igate = ipa_from_sa(&gate);
```

BIRD BUGFIXES

```
f8bcb037 - Netlink: Allow RTA_VIA even without MPLS support
--- a/sysdep/linux/netlink.c
+++ b/sysdep/linux/netlink.c
@@ -485,7 +485,6 @@ static inline ip_addr rta_get_ipa(struct rtattr *a
    )
      return ipa_from_ip6(rta_get_ip6(a));
 }

-#ifdef HAVE_MPLS_KERNEL
 static inline ip_addr rta_get_via(struct rtattr *a)
 {
   struct rtvia *v = RTA_DATA(a);
@@ -496,6 +495,7 @@ static inline ip_addr rta_get_via(struct rtattr *a
    )

+#ifdef HAVE_MPLS_KERNEL
 static u32 rta_mpls_stack[MPLS_MAX_LABEL_STACK];
```

BIRD BUGFIXES

```
d61505b0 - BSD: IPv4 over IPv6 nexthop support on FreeBSD
f8bcb037 - Netlink: Allow RTA_VIA even without MPLS support
```

BIRD BUGFIXES

```
d61505b0 - BSD: IPv4 over IPv6 nexthop support on FreeBSD
f8bcb037 - Netlink: Allow RTA_VIA even without MPLS support
```

BIRD is happy:

```
bird> show route
Table master4:
172.18.0.0/18        unicast [static1 13:39:58.386] * (200)
        via 2a12:6e40:1:1::1 on vtnet1
172.16.0.0/18        unicast [kernel1 13:40:18.040] * (10)
        via 2a12:6e40:1:1::1 on vtnet1
```

BIRD BUGFIXES

```
d61505b0 - BSD: IPv4 over IPv6 nexthop support on FreeBSD
f8bcb037 - Netlink: Allow RTA_VIA even without MPLS support
```

BIRD is happy:

```
bird> show route
Table master4:
172.18.0.0/18         unicast [static1 13:39:58.386] * (200)
        via 2a12:6e40:1:1::1 on vtnet1
172.16.0.0/18         unicast [kernel1 13:40:18.040] * (10)
        via 2a12:6e40:1:1::1 on vtnet1
```

Kernel is happy:

```
$ netstat -nr
...
172.16.0.0/18      2a12:6e40:1:1::1   UGS       vtnet1
172.18.0.0/18      2a12:6e40:1:1::1   UG        vtnet1
```

TLDR:

- Support for pfsync over IPv6 is available on FreeBSD 14.0
- Support for IPv4 over IPv6 nexthop with rtsock fixed on BIRD 2.13
- Support for IPv4 over IPv6 nexthop with Netlink fixed on BIRD 2.13.1
- FreeBSD users can finally enjoy a world with less IPv4

IT ALL COMES TO AN END...

Thank You

BACKUP

WHAT IS PFSYNC?

- pf is one of FreeBSD's built-in firewalls
- pfsync provides the syncing of the state table between two hosts
- When used together with CARP (FreeBSD's equivalent of VRRP), allows for building fully redundant firewall clusters