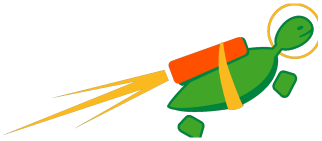


VRF Tutorial

David Ahern — Cumulus Networks

Netdev 1.1, February 2016

Agenda

VRF Overview

Use Case: Management VRF

Use Case: VLANs and VRFs

Use Case: MPLS and VRFs

Other

Virtual Routing and Forwarding (VRF)

Multiple Routing Tables

- Network interfaces correlate to routing table

Linux Implementation

- VRF represented as an L3 Master netdevice
- Interfaces are enslaved to VRF device

Applications

- Bind sockets to VRF device or cmsg + oif

L3-only concept

- Impacts only IPv4 and IPv6 route lookups

Feature by Kernel Version

IPv4 support - v4.3

IPv6 support - v4.4

“VRF-global” sockets - v4.5

Need iproute2 version that correlates to kernel version

- Or just use top of tree

Advantages of a netdevice Model

netfilter, tc rules on VRF device

tcpdump / packet capture on VRF device

IP on VRF device - VRF-local loopback

Nesting of VRFs (L3) in a namespace (L1)

Applications can use existing APIs

- cmsg & oif, SO_BINDTODEVICE, IP_PKTINFO

Existing framework for serviceability and debugging

- e.g., list VRFs using ’ip link show type vrf’

Conceptual Basics

Create VRF device with table id

- ip link add vrf-red type vrf table 123

Setup FIB rules pointing to device

- ip {-6} ru add pref 200 {i,o}if vrf-$name table $tbid

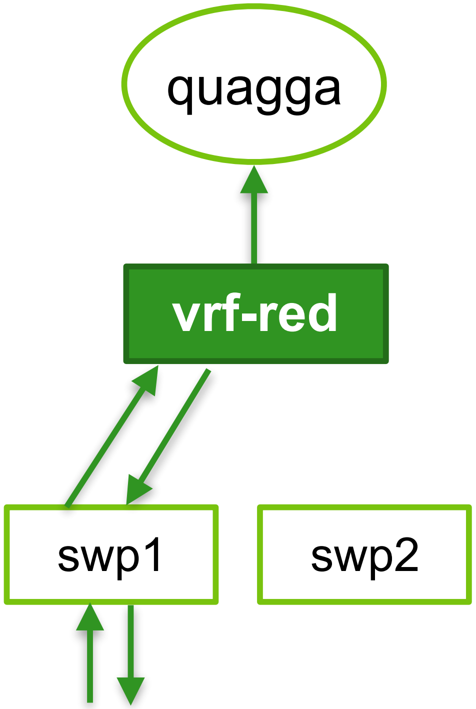
Enslave interfaces

- ip link set swp1 master vrf-red

Bind socket to VRF device

- Connects socket to L3 domain

Send / receive packets



VRF Create

‘vrf_create <name> table <id>’ means:

```
ip link add vrf-${name} type vrf table ${tbid}
echo "${tbid}  vrf-${name}" > /etc/iproute2/rt_tables.d/vrf-${name}.conf
ip ru add pref 200 oif vrf-${name} table ${tbid}
ip ru add pref 200 iif vrf-${name} table ${tbid}
ip -6 ru add pref 200 oif vrf-${name} table ${tbid}
ip -6 ru add pref 200 iif vrf-${name} table ${tbid}
ip route add table ${tbid} unreachable default
ip -6 route add table ${tbid} unreachable default
ip link set dev vrf-${name} up
```

Use Cases

Management VRF

VLANs and VRFs

MPLS and VRFs

Use Case: Management VRF

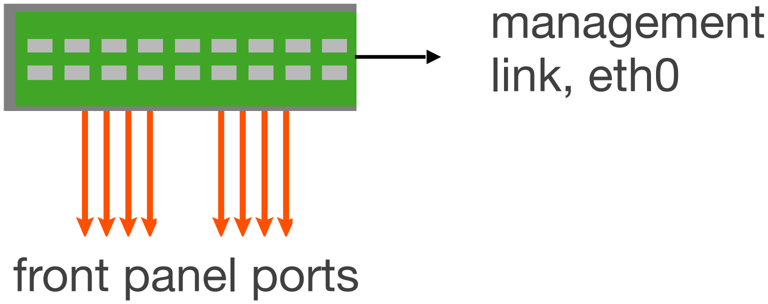
Separation of Management Traffic from Data Plane

Management Interface

- Table 252

Front Panel Ports

- Main Table



Use Case: Management VRF

Create VRF device

- vrf-create mgmt table 252

Enslave management interface

- ip link set dev eth0 master vrf-mgmt

List routes in management table

- ip route ls table vrf-mgmt

Use Case: Management VRF

Commands by default use main table == front panel ports

- ping, traceroute, etc

Management apps

- need to be told to use mgmt table
  - e.g, ping -I vrf-mgmt <ip>
- SO_BINDTODEVICE, somark, LD_PRELOAD that sets device binding
  - l3mdev cgroup solves this problem more elegantly

Live Demo

Use Cases

Management VRF

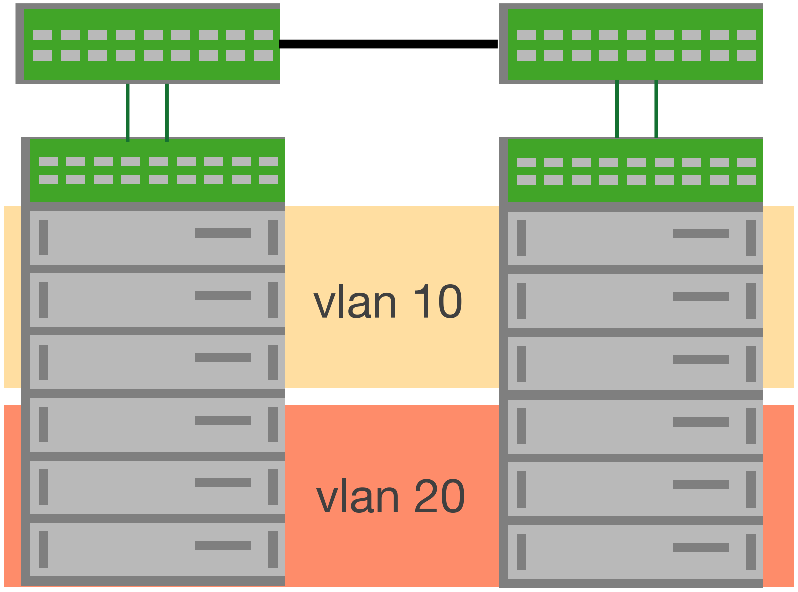
VLANs and VRFs

MPLS and VRFs

VRF with VLANs

Hosts separated by VLANs

Shared networking infrastructure



Use Case: VRFs and VLANs

Demonstration using simplified deployment model

2 groups of hosts

g1h{1-4} - isolated on VLAN 10

g2h{1-4} - isolated on VLAN 20

4 switches

- hosts connect to switches s1 and s2
- routing provided by switches r1and r2
  - VRF 10 and VRF 20

Live Demo

ASCII Art diagram of demo

s1 config

VLAN 10 subinterfaces on swp1, swp2, swp5

bridge br1 with swp1.10, swp2.10 and gateway address for hosts g1h1, g1h2

VLAN 20 subinterfaces on swp3, swp4, swp6

bridge br2 with swp3.20, swp4.20 and gateway address for hosts g2h1, g2h2

swp5.10 and swp6.20 have addresses to route packets to r1 for respective vlans

```
  10.1.1.1    10.1.1.2       10.2.1.1    10.2.1.2
  +======+    +======+       +======+    +======+
  | g1h1 |    | g1h2 |       | g2h1 |    | g2h2 |
  +======+    +======+       +======+    +======+
      |           |             |           |
  +=============================================+
  |   1           2             3           4   |
  |  1.10        2.10          3.20        4.20 |
  |         br1                     br2       | br1: 10.1.1.254
  |                      s1                     | br2: 10.2.1.254
  |         5.10                    6.20        |
  |          5                       6          |
  +=============================================+
          .2 |                       | .2
10.10.1.0/28 |                       |    10.20.1.0/28
          .1 |                       | .1
          +=============================+
  VRF 10  |  2                       3  |   VRF 20
          | 2.10       r1          3.20 |
          |                             |
```

ASCII Art diagram of demo

r1 config

VLAN 10 subinterfaces on swp1, swp2

VLAN 20 subinterfaces on swp1, swp3

All VLAN interfaces have addresses for routing packets

r2 config is similar to r1

```
             .2 |                       | .2
10.10.1.0/28    |                       |    10.20.1.0/28
   VLAN 10   .1 |                       | .1   VLAN 20
             +=============================+
     VRF 10  |  2                       3  |   VRF 20
             | 2.10       r1       3.20 |
             |                             |
             |         1.10 1.20           |
             |             1               |
             +=============================+
                        .1 |
           10.99.1.0/30    | VLAN trunk, 10 and 20
                        .2 |
             +=============================+
             |             1               |
             |         1.10 1.20           |
             |                             |
             | 2.10       r2       3.20 |
    VRF 10   |  2                       3  |    VRF 20
             +=============================+
   VLAN 10   .2 |                       | .2   VLAN 20
10.10.2.0/28    |                       |    10.20.2.0/28
             .3 |                       | .3
```

ASCII Art diagram of demo

`s2 config mirrors s1`

```
          |                             |
          | 2.10       r2          3.20 |
 VRF 10   |  2                       3  |    VRF 20
          +=============================+
          .2 |                       | .2
10.10.2.0/28 |                       |    10.20.2.0/28
          .3 |                       | .3
  +=============================================+
  |          5                       6          |
  |         5.10                    6.20        |
  |                      s2                     |
  |         br1                      br2        | br1: 10.1.2.254
  |  1.10        2.10          3.20        4.20 | br2: 10.2.2.254
  |   1           2              3           4  |
  +=============================================+
      |           |              |           |
  +======+    +======+       +======+    +======+
  | g1h3 |    | g1h4 |       | g2h3 |    | g2h4 |
  +======+    +======+       +======+    +======+
  10.1.2.3    10.1.2.4       10.2.2.3    10.2.2.4
```

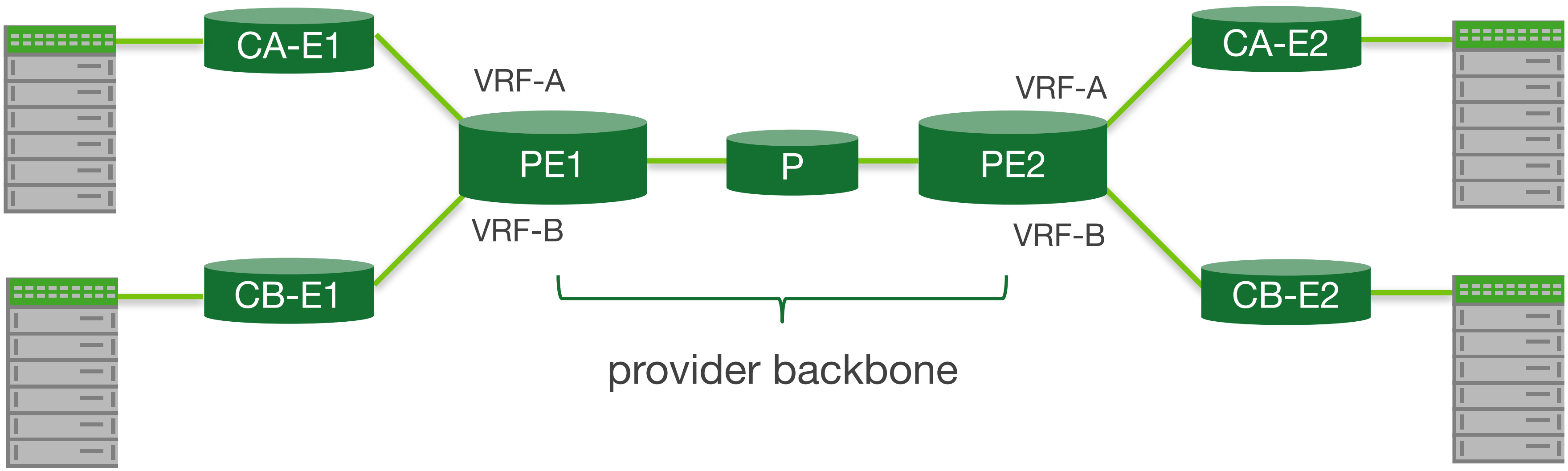
Use Cases

Management VRF

VLANs and VRFs

MPLS and VRFs

VRF with MPLS



Use Case: VRFs and MPLS

Demonstration using simplified deployment model

2 customers with hosts separated by common pe1-p-pe2 backbone

backbone uses MPLS

hosts connected to edge routers

Edge routers connected to provider edge

- VRFs at provider edge provide the separation

Live Demo

ASCII Art diagram of demo

typical address assignment for hosts, cNhM

Edge switch is the default gateway

```
         88.2.1.3    88.2.1.4         88.2.1.3    88.2.1.4
         +======+    +======+         +======+    +======+
         | c1h3 |    | c1h4 |         | c2h3 |    | c2h4 |
         |    1 |    |    1 |         |    1 |    |    1 |
         +======+    +======+         +======+    +======+
             |            |                |           |
88.2.1.0/24  +---+    +---+                +--+     +--+ 88.2.1.0/24
                 |    |                       |     |
         +===============+            +===============+
         |       2    3  |            |       2     3 |
         |         br0   |88.2.1.254  |         br0   | 88.2.1.254
         |               |            |               |
         |  1   c1e2     |            |  1   c2e2     |
         +===============+            +===============+
         .2 |                            |.2
            | 3.1.1.0/30      3.2.1.0/30 |
            +--------+       +-----------+
                     |       |
```

ASCII Art diagram of demo

edge switches attached to provider backbone at both ends

pe1 and pe2 are LERs. VRFs separate customer traffic coming into and out of PEs

p is an LSR

```
                   .1 |       | .1
          VRF 10  +==================+   VRF 20
push label 102 v  |    2        3    |  v  push label 202
                  |        pe2       |
 pop label 111 ^  |         1        |  ^  pop label 211
                  +==================+
                         .6 |
              2.1.1.4/30    |
                         .5 |
 pop label 102 v  +==================+  v pop label 202
push label 112    |        swp2      |    push label 212
                  |               p  |
 pop label 101 ^  |        swp1      |  ^ pop label 201
push label 111    +==================+    push label 211
                         .2 |
              2.1.1.0/30    |
                         .1 |
                  +==================+
 pop label 112 v  |         3        |  v pop label 212
                  |        pe1       |
push label 101 ^  |    1        2    |  ^ push label 201
          VRF 10  +==================+  VRF 20
                    .2 |       | .2
```

ASCII Art diagram of demo

Customer hosts on the other side of the PE1-P-PE2 backbone

```
            1.1.1.0/30    |       |    1.1.1.0/30
               +----------+       +-----------+
               |                              |
            .1 |                              | .1
            +================+             +================+
            |  1   c1e1      |             |  1   c2e1      |
            |                |             |                |
            |         br0    | 88.1.1.254  |         br0    |
88.1.1.254
            |       2    3   |             |       2      3 |
            +================+             +================+
                    |    |                         |     |
   88.1.1.0/24  +---+    +---+                  +--+     +--+
88.1.1.0/24
                |            |                  |           |
            +======+    +======+           +======+    +======+
            |    1 |    |    1 |           |    1 |    |    1 |
            | c1h1 |    | c1h2 |           | c2h1 |    | c2h2 |
            +======+    +======+           +======+    +======+
            88.1.1.1    88.1.1.2           88.1.1.1    88.1.1.2
```

Inter-VRF Routing

Routing Between VRFs

Explicit route in a table

- ip route add table vrf-red 1.1.1.0/24 dev eth2
- eth2 is in alternate VRF

Full lookup in VRF table

- ip route add table vrf-red 1.1.1.0/24 dev vrf-green

Misc

cumulus®

Link down on VRF device stops all routing in that VRF

- Similar to a bridge and ports enslaved to it

Address on VRF device

- VRF local loopback address

Overlapping Addresses

- design allows same address/prefix in different VRF domains cumulus®

Unleashing the Power of Open Networking

Thank You!