1 Route distinguisher in mpls networks

Route distinguishers (or RD) in an mpls networks have been created to provide the chance to discriminate on a PE the same prefix used in two different vrf. From RFC 2547 we can find the following definition:

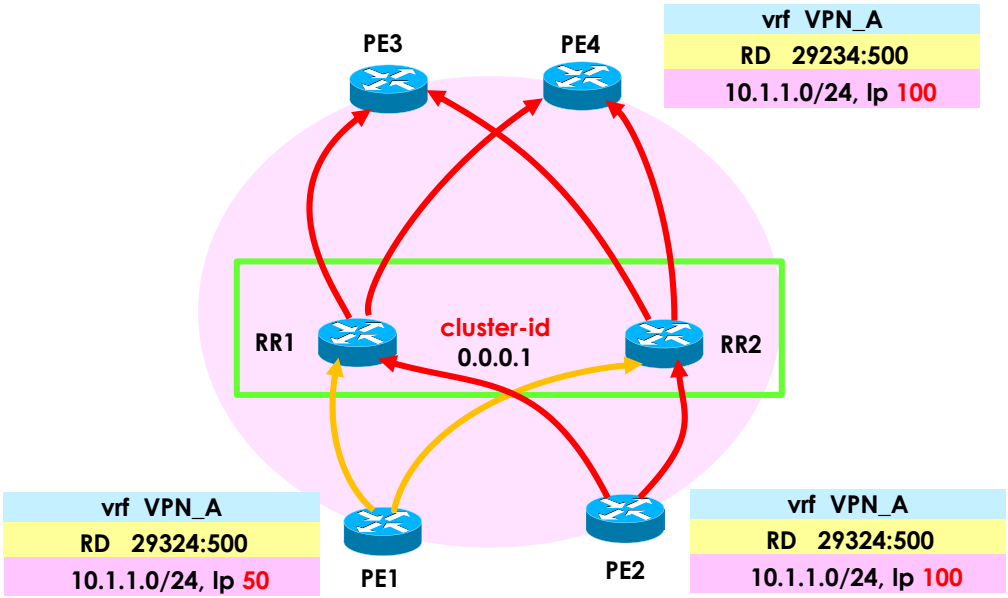
> " The BGP Multiprotocol Extensions [3] allow BGP to carry routes from multiple "address families". We introduce the notion of the "VPN-IPv4 address family". A VPN-IPv4 address is a 12-byte quantity, beginning with an **8-byte "Route Distinguisher (RD)"** and ending with a 4-byte IPv4 address. If two VPNs use the same IPv4 address prefix, the PEs translate these into unique VPN-IPv4 address prefixes. This ensures that if the same address is used in two different VPNs, it is possible to install two completely different routes to that address, one for each VPN."

It is an accepted best practice to write the RD as follows:

<BGP AS> : <value>

… where the first value is the bgp as number used by the mpls network, and the value is a 32bit number. There are no rules in chosing which RD value to assign to each vrf, except for the obvious constraint that on the same PE **two different vrf MUST NOT have the same rd value**. In a real mpls network, some people choose a unique RD value and configure it on all PEs. What changes ? what are the available approaches ? Beware that a vpnv4 route is the sum of:

vpnv4 route → RD + ipv4 prefix



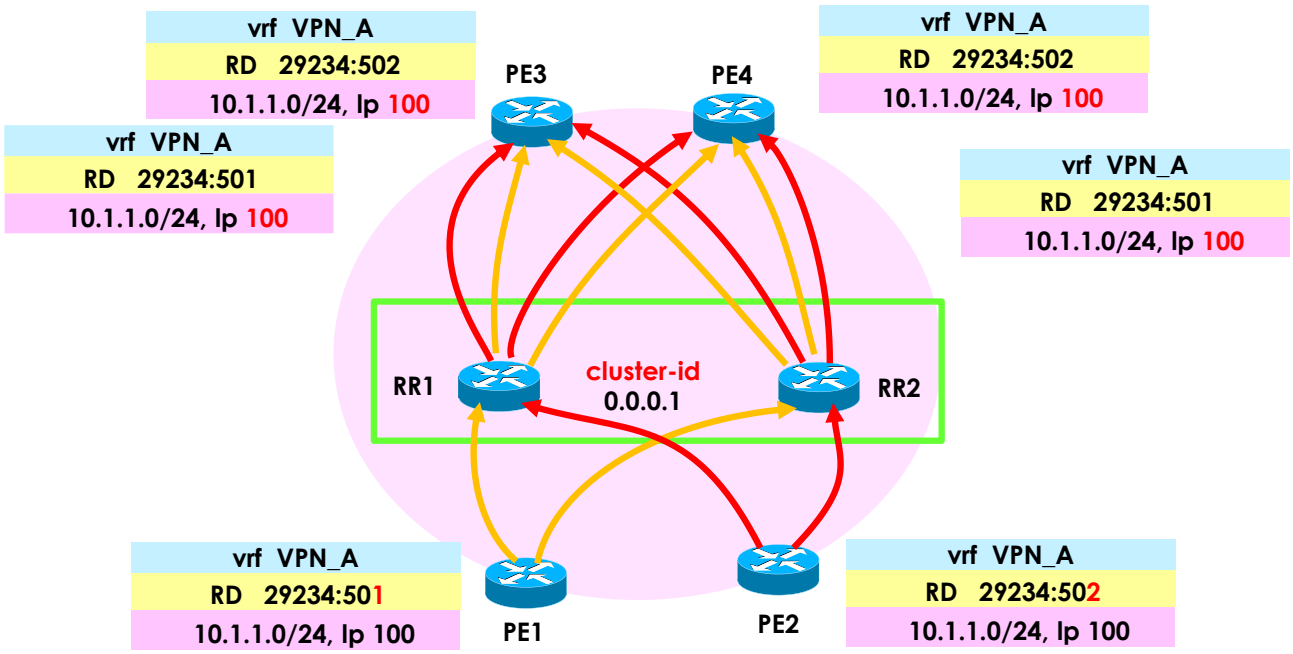
Consider the above example: in case the same RD is configured on all PEs for the same vrf, route-reflectors in the network will forward **only the best bgp route**. This could lead to both RR forwarding the SAME vpnv4 route, for example that announced by PE2. In this case PE3 and PE4 will have only one next-hop (PE2) in their routing table for subnet 10.1.1.0/24. This could happen even in case the local-preference is the same, the tie-braker being the igp cost toward the announcing PE.

This approach has two negative consequences:

- you can't configure any load-balancing on PE3 and PE4
- convergence time will be higher, since the two RR will need to announce PE1's route when PE2's route is no more valid

The only advantage is potential memory saving, since PE3 and PE4 will receive less routes.

This is what happens in case RD values are set different, even if the vrf is the same:



PE3 and PE4 see both the routes from PE1 and PE2. In case of problems on PE2, when the route is withdrawn on PE3 and PE4 they already have locally and ready the backup route announced by PE1, this will improve convergence time. Moreover, ibgp load balancing could be configured:

```
Router(config)# router bgp 40000
 Router(config-router)# address-family ipv4 vrf RED
 Router(config-router-af)# maximum-paths eibgp 4      ←this balances also eBgp routes
 Router(config-router-af)# end
```

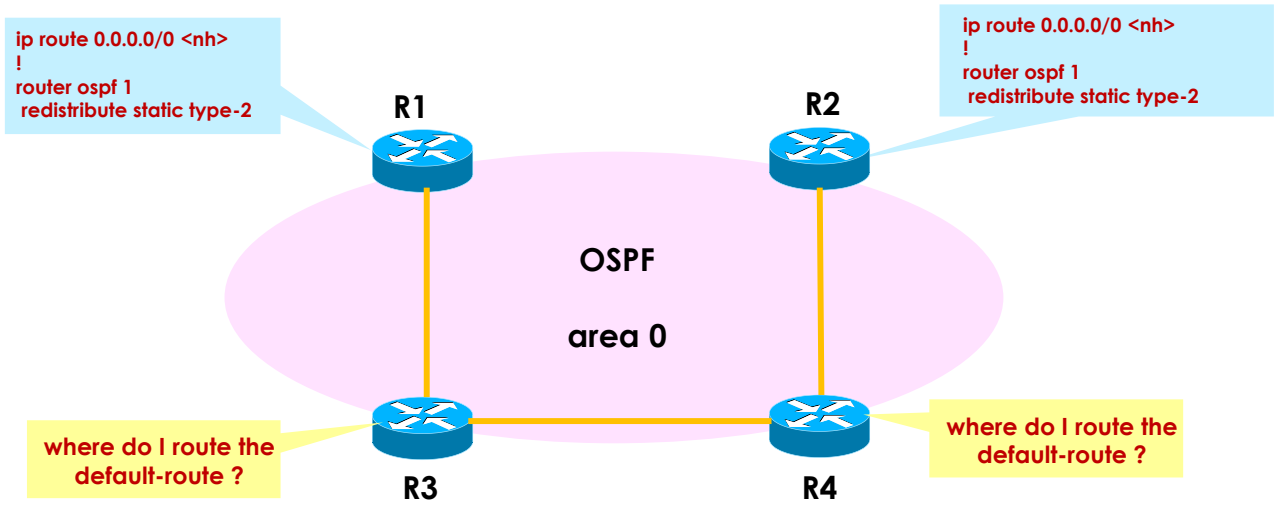
The above command "Configures the number of equal-cost or **unequal-cost** routes that BGP will install in the routing table.", unequal depending of course on bgp attributes, some of the most important bgp parameters (local-preference, as-path, metric, …) must be the same for all the available routes. The igp cost to the next hop can be different, and you can balance between two (or more) PEs with no risks, since traffic is label switched and not routed hop by hop.
Beware that having SO MANY rd values available, it would make sense for troubleshooting purposes to give a meaning to those values, for example something like the following:

```
<3 digits node id> <4 digits vpn id>
```

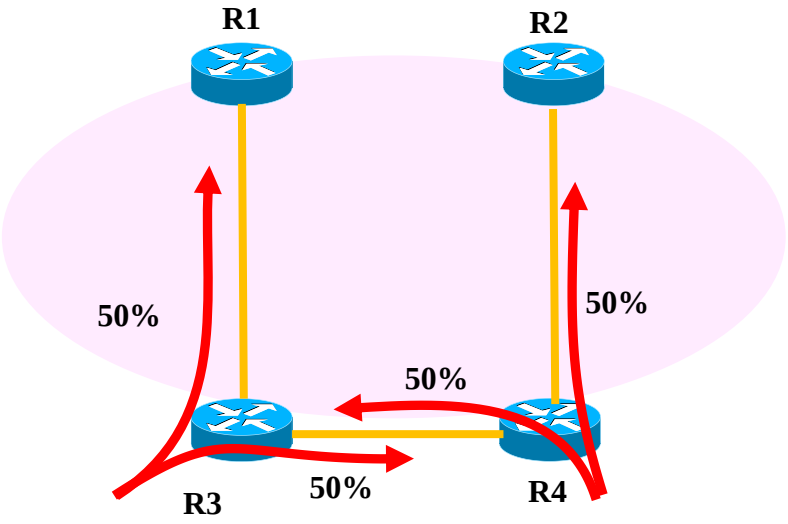
1.1 Unequal cost load balancing

Let's consider unequal cost balancing with 'legacy' IGP link-state routing protocols. Would it be possible ? with EIGRP it is, even though you should take care about what happens, and how you configure it. Why ? Consider this very simple topology with OSPF routing protocol. From RFC 2328:

"Processing of Type 2 external metrics is simpler. The AS boundary router advertising the smallest external metric is chosen, regardless of the internal distance to the AS boundary router. Suppose in our example both Router RT5 and Router RT7were advertising Type 2 external routes. Then all traffic destined for Network N12 would be forwarded to Router RT7, since 2 < 8. When several equal-cost Type 2 routes exist, **the internal distance to the advertising routers is used to break the tie**.
Both Type 1 and Type 2 external metrics can be present in the AS at the same time. In that event, Type 1 external metrics always take precedence."



In case the default-routes are exported with the SAME metric, the internal distance to the advertising routers must be considered. This means that R3 will route traffic toward R1, and R4 will route traffic toward R2 (we suppose that all link's cost is the same of course). What would happen in case R3 and R4 perform unequal cost load balancing ? what if they balance traffic toward the two advertising routers R1 and R2 ?



In this case, supposing that balancing works perfectly and that (as routers usually do) per-flow load balancing is in place, 25% of the total traffic would be dropped due to a loop between R3 and R4. This is why ospf stands that in case two routers are advertising the same external type-2 prefix, you need to prefer the nearest advertising router to break the tie. Forwarding in an mpls network works differently, since you label switch hop by hop toward a next-hop that has been decided by the ingress router.