Analysis of VPN Routes Control in Shared BGP Session

draft-wang-idr-vpn-routes-control-analysis

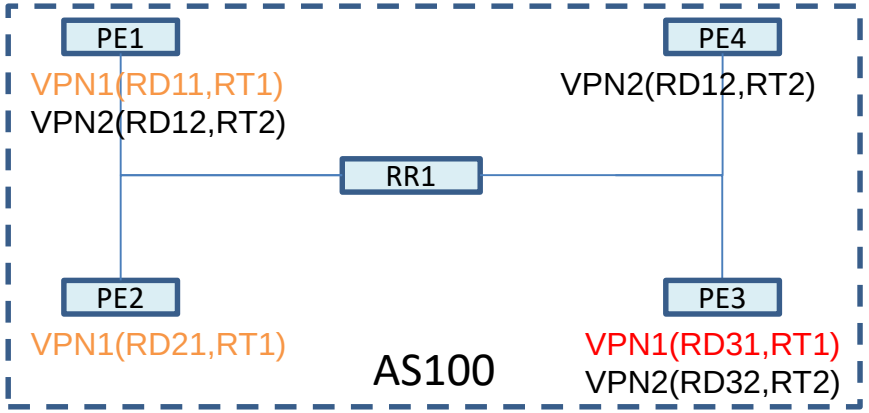
Aijun Wang (China Telecom)
Wei Wang (China Telecom)
Gyan Mishra (Verizon)
Haibo Wang (Huawei)
Shunwan Zhuang (Huawei)
Jie Dong (Huawei)

IETF 110, March. 2021

Motivation of This Draft

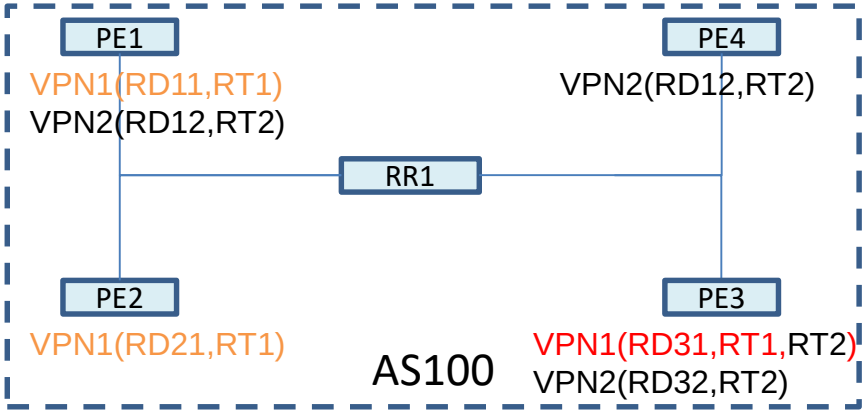
- ✓ Describes the scenarios that need to control VPN routes in one shared BGP session
- ✓ Reaches consensus on the problem analysis and solution requirements
- ✓ Establishes the base for the afterward updated solution

Start from the simplest scenario(intra-AS)



CASE 1

① RD is allocated per VPN/per PE
② Only one RT is associated with such VPN routes
③ PE3 send excessive VPN routes with RT1

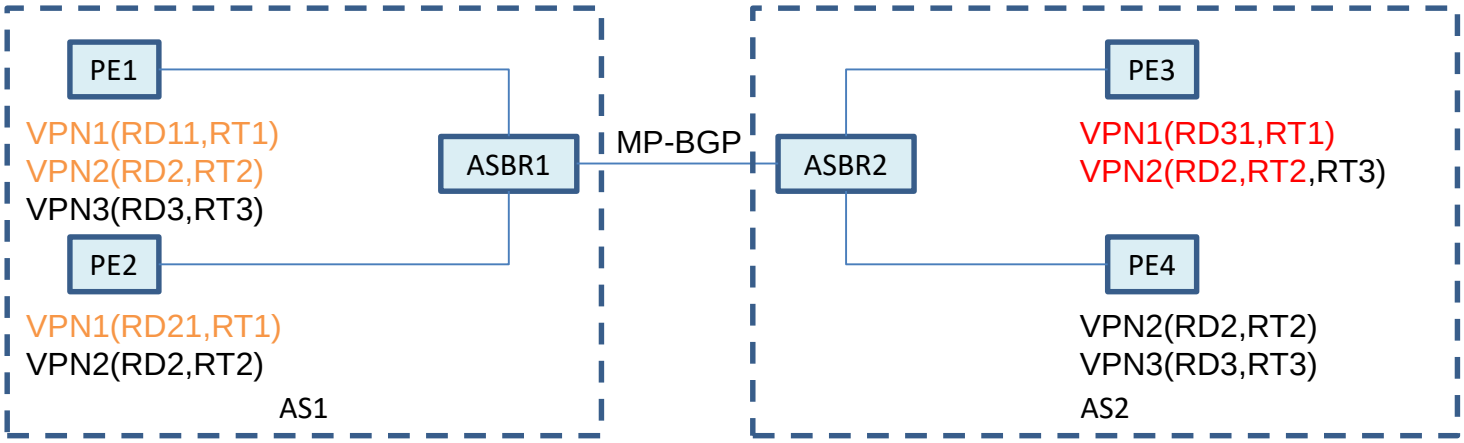


CASE 2

① RD is allocated per VPN/per PE
② Multiple RTs are associated with such VPN routes, and be imported into different VRFs in other devices(PE1)
③ PE3 send excessive VPN routes with RT1.

1. The BGP session between RR and PE is shared for the VPN routes advertisements.
2. The excessive VPN routes from one VRF can also influence the PE's performance for other VRFs.
3. Reason for the excess VPN routes may be the followings(as Jim Uttaro mentioned)
   a) Multiple CEs connect to PE3 advertising routes simultaneously causing a wave of routes,
   b) Redistribution from VRF to VRF, Or from GRT to VRF, or Other error configuration etc.
4. PE/RR should have some mechanisms to identify and control the advertisement of specified excessive VPN routes.

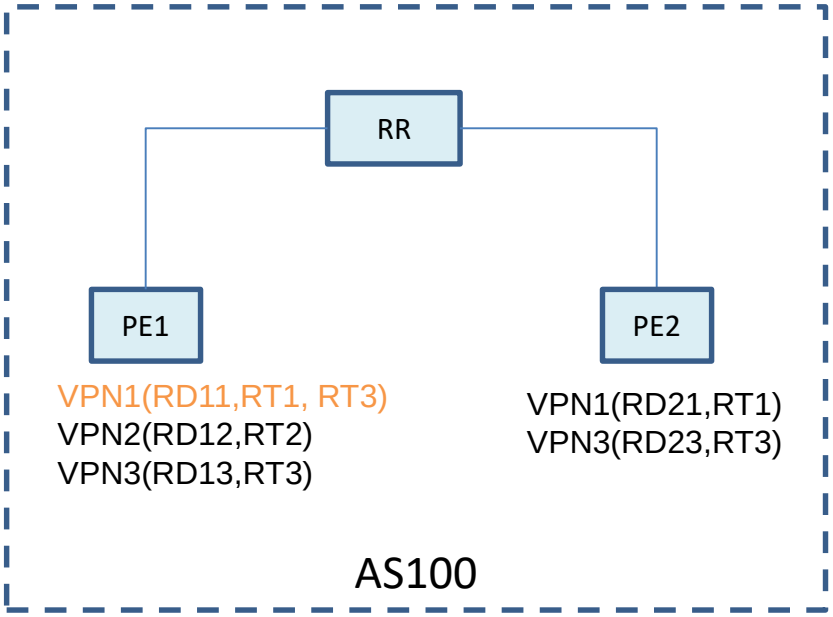
Inter-AS Scenario



1. For inter-AS scenario(Option B/AB/C), BGP session is shared between PE/ASBR and ASBR1/ASBR2
2. The possibility of excessive VPN routes advertisement is same as that in the intra-AS.
3. Overwhelming VPN routes in one VRF can certainly influence the control/forward behavior of the PE for other VRFs.
   a) RD may be allocated per VRF per PE, as that in intra-AS(VPN1 in above figure).
   b) RD may also be allocated per VRF only. That is to say, same VPN on different PEs has same RD. (VPN2、VPN3 in above figure)
4. The ASBR/PE should have some capabilities/mechanism to distinguish and control the advertisement of specified excessive VPN routes.

VPN routes share on one PE Scenario

1. Excepts the previous VPN routes sharing on one BGP session, such VPN routes may be shared among different VRFs on one device, as illustrated in figure.
2. If only some of the VRF has reach its VPN route limit because of the shared VPN routes, the PE device should have the capability to control the distribution of such VPN routes from its global VPN table.
3. Only when all of the shared VRFs don't want such VPN routes, the PE can then send out the control message to its upstream peer.



Solution Requirements

The potential solutions should meet the following requirements:

a) Control message for the specified VPN routes should be triggered automatically upon the excessive VPN routes reach its limit.

b) Control message should be sent only out the device:

- ✓ For PE: when all the VRFs on it don't want to process it
- ✓ For RR: when all its BGP clients don't want to process it
- ✓ For ASBR: when all its BGP peers within one AS don’t want to process it
- ✓ Or for all of them: the process of such excessive routes has exceed its own capability.

c) The trigger and removal of such control message should avoid the possible flapping of excessive VPN routes advertisement.

Further Action

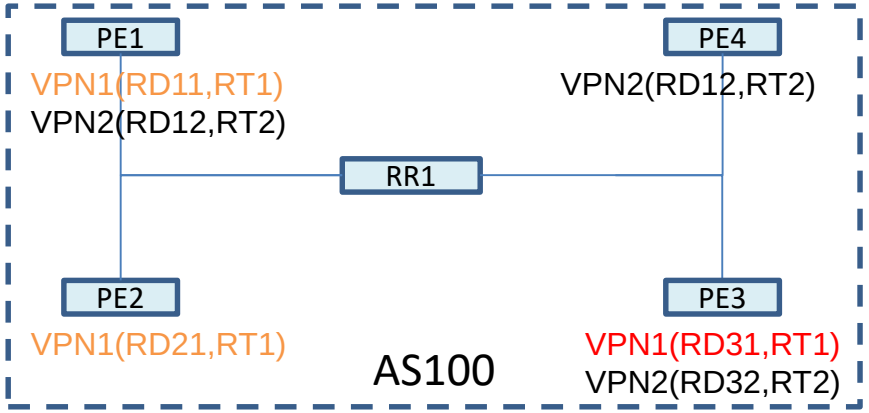
- Comments?
- Is this clear to describe the problem and solution requirement
- If so, then start the design and update of the solution draft?

wangaj3@chinatelecom.cn
wangw36@chinatelecom.cn
gyan.s.mishra@verizon.com
rainsword.wang@huawei.com
zhuangshunwan@huawei.com
jie.dong@huawei.com
IETF110

Potential/Proposed Solution

- If we have consensus on such direction, will update the draft and discuss it in detail later

Start from the simplest scenario(intra-AS)



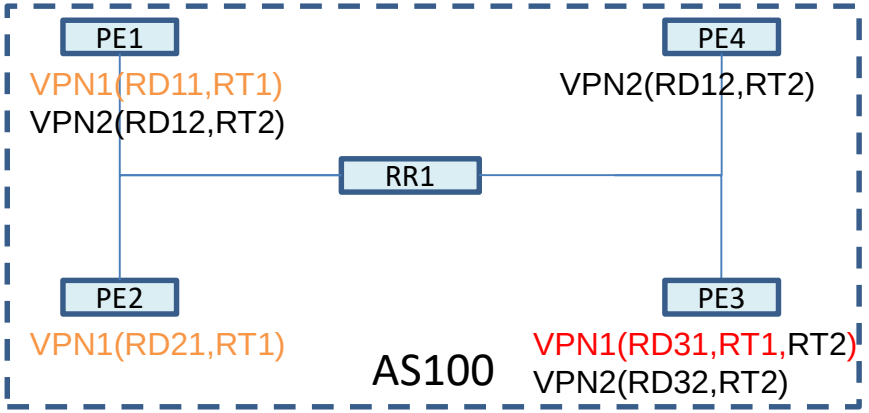
CASE 1

① Once PE1 detects the VPN1 VRF is overflowed , and:
- ✓ The excessive VPN routes has RD31,
- ✓ No other VRFs on it import the VPN routes with RT 1, which is the only RT that attached with such excessive VPN routes

✓ PE1 triggers the RD-ORF message to RR1(RD field is set to RD31)

✓ RR1 withdraws and stops to advertise such excessive VPN routes to PE1

✓ Other VPN services on PE1 will keep to service.



CASE 2

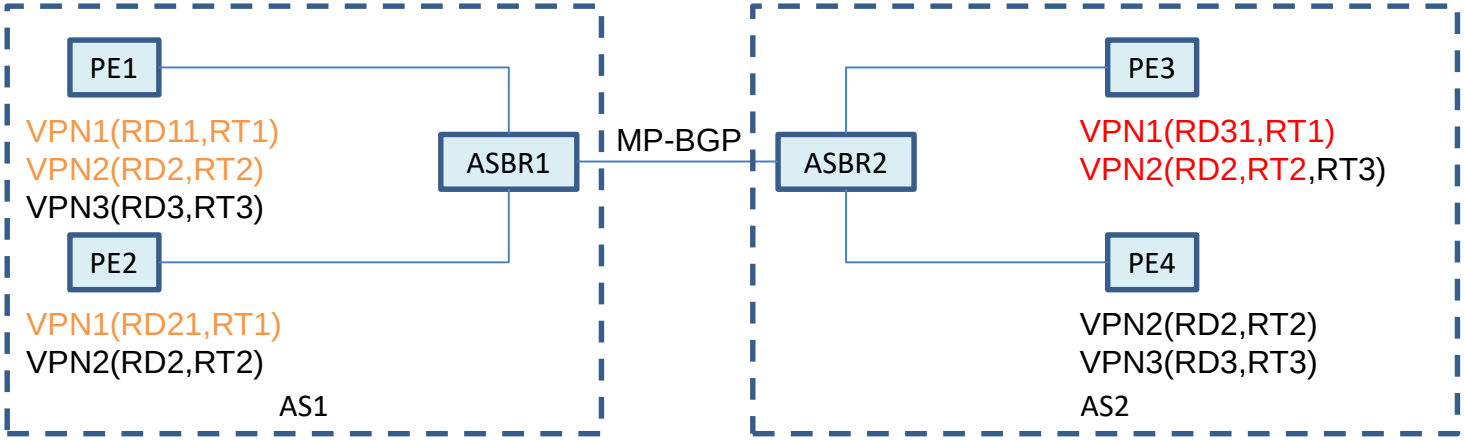
① Once PE1 detects the VPN1 VRF is overflowed , and:
- ✓ The excessive VPN routes has RD31,
- ✓ There is other VRF on it import the VPN routes with RT 2, which is the RT that attached with such excessive VPN routes

✓ PE1 will discard such excessive VPN routes on it and hold the RD-ORF message to RR1

✓ RR1 still send such excessive VPN routes to PE1 until it reach its process capabilities.

✓ Other VPN services on PE1 will keep to service.

Inter-AS Scenario

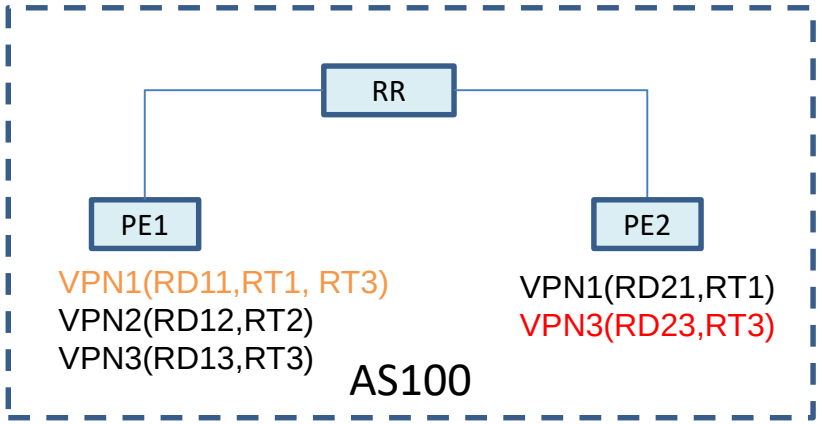


1. If the excessive VPN routes are from VPN1 or VPN2 on PE3, once the PE1 detects the overflow of VPN1 on it:
   - ✓ If there is no other VPNs on it import the RT that the excessive VPN routes attached.
   - ✓ PE1 will trigger the RD-ORF(RD field is set to RD31 or RD2) message to ASBR1

2. When ASBR1 receives such RD-ORF message, it check:
   - ✓ If all its downstream peers sent the same message, or the process of excessive VPN routes have exceed its capabilities, it will send such message to upstream peer(ASBR2)
   - ✓ Or else, it will filter the excessive VPN routes on its side, on behalf of the trigger device(PE1)
3. For example, in above figure:
   - ✓ If PE1 and PE2 all sent the RD-ORF(with RD field set to RD31) message, the ASBR1 can send out the RD-ORF(with RD field set to RD31) message then.
   - ✓ If only PE1 sent the RD-ORF(with RD field set to RD2) message, ASBR1 will filter the excessive VPN routes to PE1. PE2 can still receive such routes.

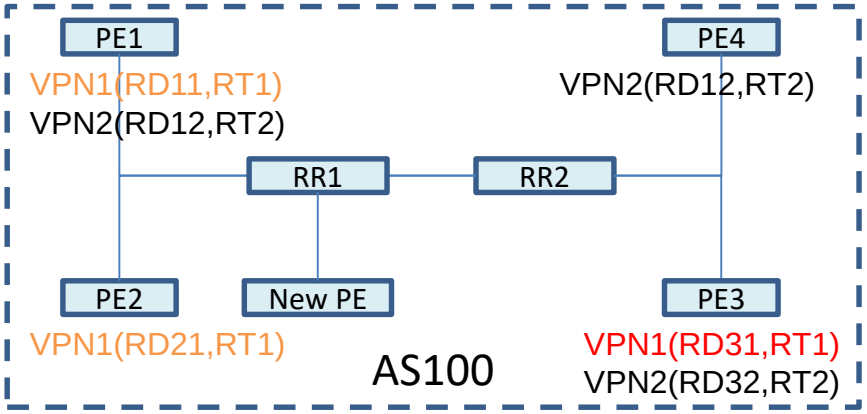
RD-ORF Trigger condition on PE and RR

1. If PE receives the excessive VPN routes, and there are several VRFs on the PE import such routes:
   - ✓ Then PE should not trigger the RD-ORF messages until all such VRFs declare that they don't want such routes, or such process exceed its capabilities. Or else:
   - ✓ The PE should filter such routes on its side.



Behavior on receiving PE

1. If RR receives the same RD-ORF message from all its clients(does not include the peer, for example RR2), or process such excess routes exceed its capabilities:
   - ✓ RR can then send out the RD-ORF message to its upstream peer(RR2 in this graph).
2. If one new PE is attached under the RR, then RR should revoke the sent RD-ORF message to its upstream peer：
   - ✓ New PE will receive the excessive VPN routes, until it reaches its maximum threshold and send out RD-ORF



Behavior on RR

Other Solutions/Consideration

- Welcome feedbacks/comments/other solutions