BGP Signaled IPsec Tunnel Configuration

Hu Jun, Nokia

3/5/2019

Problem

In some networks, there is need to encrypt traffic between routers, which is typically done by putting traffic into IPsec tunnel; however when the number of router is big, it is difficult to provision and manage large number of mesh IPsec tunnels on all routers, specially when a simple hub-and-spoke topology can’t use used;

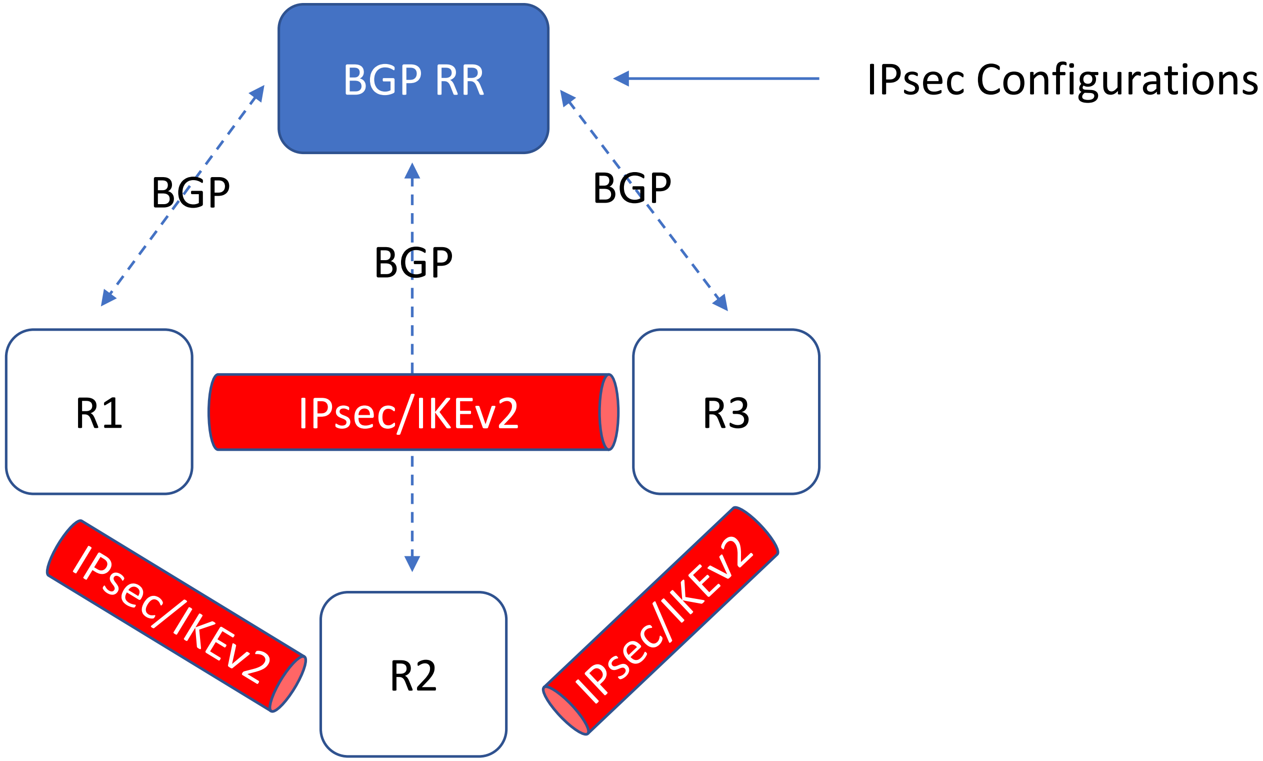
A more efficient way to provision IPsec tunnel is needed for such cases;

Design Considerations

1. Not trying to be a cure for everything, just to address IPsec config provision problem
2. The solution shouldn’t trade off security
3. Avoid reinventing wheel, reuse existing protocol wherever possible, and keep changes minimal

Solution

- This draft defines a method of using BGP to signal IPsec tunnel configuration along with NLRI, it uses and extends tunnel encapsulation attribute as specified in [I-D.ietf-idr-tunnel-encaps] for IPsec tunnel.
- BGP is only used to signal certain IPsec configuration, the IPsec tunnel is still created via IKEv2 between routers after the configuration is learned via BGP UPDATES.



BGP Tunnel Encapsulation Attribute Extensions for IPsec

This document extends tunnel encapsulation attribute specified in [I-D.ietf-idr-tunnel-encaps] by introducing following changes:

- A tunnel type for IPsec tunnel: ESP tunnel mode (AH tunnel mode is not included in this document). Existing type 4 (IPsec in Tunnel-mode) in IANA "BGP Tunnel Encapsulation Attribute Tunnel Types" registry could be reused
- A new sub-TLV for public routing instance: where IPsec packet is forwarded in, which could be different from payload packet
- A new sub-TLV for remote address prefix: remote traffic selector (from receiver POV)
  - Another way to do this is to use recursive lookup, but need more updates
- A new sub-TLV for local address prefix: local traffic selector (from receiver POV)

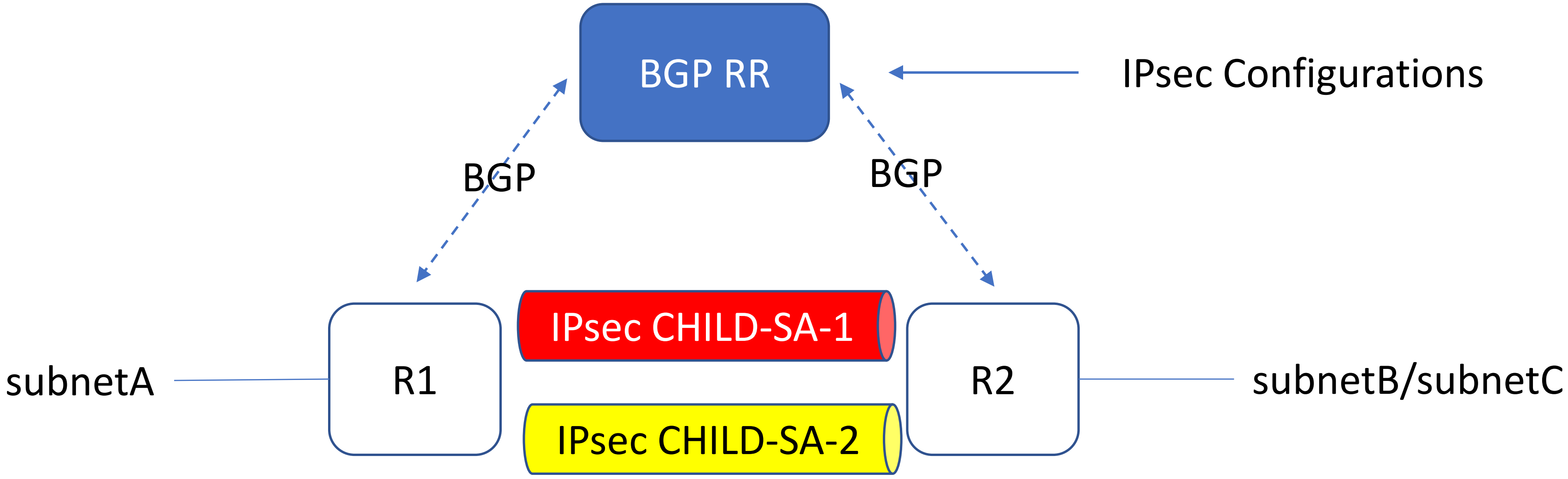
It also reuses following existing sub-TLV:

- Remote Endpoint: IPsec tunnel endpoint address
- Color: IPsec configuration attributes like ESP transform; the meaning of this sub-TLV is local to the administrative domain
- Embedded Label Handling: see section 4 of draft for detail

Operation Example

Requirements:

- Traffic between subnetA - subnetB: ESP tunnel, AES-CBC-256 with SHA-384, mapping to color red
- Traffic between subnetA - subnetC: ESP tunnel, null encryption with only integrity protection, SHA-256, mapping to color yellow



Operation Example (Cont.)

1. Both R1 and R2 are provisioned with PKI key and certificate from same CA.
2. R1 advertise subnetA in BGP UPDATE, which has a tunnel encapsulation attribute that contains IPsec TLVs:
   - TLV-1: Remote Endpoint sub-TLV R1TunnelAddr, color sub-TLV red and subnetB in Local Prefix sub-TLV.
   - TLV-2: Remote Endpoint sub-TLV R1TunnelAddr, color sub-TLV yellow and subnetC in Local Prefix sub-TLV.
3. R2 advertise subnetB in BGP UPDATE, which has a tunnel encapsulation attribute that contains one IPsec TLV : Remote Endpoint sub-TLV R2TunnelAddr, color sub-TLV red and subnetA in Local Prefix sub-TLV.
4. R2 advertise subnetC in BGP UPDATE, which has a tunnel encapsulation attribute that contains one IPsec TLV: Remote Endpoint sub-TLV R2TunnelAddr, color sub-TLV yellow and subnetA in Local Prefix sub-TLV.

Operation Example (Cont.)

5. R1 received a packet from subnetA destined to subnetB, since BGP UPDATE contain subnetB also contains an IPsec tunnel encapsulation attribute, there is no existing CHILD SA could be used, R1 select TLV-1 and uses IKEv2 to establish an IPsec tunnel to R2TunnelAddr, using certificate authentication, create 1st CHILD SA CHILDSA1:
   - ESP transform: AES-CBC-256 and SHA-384
   - Traffic Selector:
     - TSi: address subnetA, protocol any, port any
     - TSr: address subnetB, protocol any, port any
6. after tunnel is created, R1 and R2 could forward traffic between subnetA and subnetB over CHILDSA1
7. R1 received a packet from subnetA destined to subnetC, CHILDSA1 can't be used for this packet, R1 select TLV-2 to create 2nd CHILD SA, and given there is already an IKE SA between R1 and R2, R1 uses existing IKESA to create CHILDSA2:
   - ESP transform: Null encryption with SHA-256
   - Traffic Selector:
     - TSi: address subnetA, protocol any, port any
     - TSr: address subnetC, protocol any, port any
8. R1 and R2 could forward traffic between subnetA and subnetC over CHILDSA2

Summary & Benefits

This draft propose a method using BGP to signal IPsec configuration by extending BGP Tunnel Encapsulation attribute, while still using IKEv2 to create the tunnel.

Main benefits of this proposal are:

- Decouple IPsec configuration from key exchange and tunnel negotiation, be able to leverage best of each side; allow future development independently.
- By reusing IKEv2, a field proven, mature protocol, not sacrificing security
- Only limited extension to BGP, no change to IPsec/IKEv2
- Inter-op with router that doesn’t support this draft, which allow graceful transition

Next Step?

Support other types of encryption tunnel like DTLS?