IDR Working Group
Internet-Draft
Intended status: Informational
Expires: January 15, 2021

S. Hares
Hickory Hill Consulting
July 14, 2020

BGP IPSEC VPNs - A Solution Analysis
draft-hares-idr-bgp-ipsec-analysis-00

Abstract

This draft describes problems with IGP convergence time in some IPRAN networks that use a physical topology of grid backbones that connect rings of routers. Part of these IPRAN network topologies exist in data centers with sufficient power and interconnections, but some network equipment sits in remote sites impacted by power loss. In some geographic areas these remote sites may be subject to rolling blackouts. These rolling power blackouts could cause multiple simultaneous node and link failures. In these remote networks with blackouts, it is often critical that the IPRAN phone network re-converge quickly.

The IGP running in these networks may run in a single level of the IGP. This document seeks to briefly describe these problems to determine if the emerging IGP technologies (flexible algorithms, dynamic flooding, layers of hierarchy in IGPs) can be applied to help reduce convergence times. It also seeks to determine if the improvements of these algorithms or the IP-Fast re-route algorithms are thwarted by the failure of multiple link and nodes.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document analyzes three solutions for using a BGP based approach on SDWAN edge nodes to establish secure IPSec tunnels for the overlay routes distribution. The solutions are:

- [I-D.hujun-idr-bgp-ipsec]

- [I-D.dunbar-idr-sdwan-edge-discovery]

- [I-D.sajassi-bess-secure-evpn]

These three drafts propose an IPsec related tunnel type for an augmentation of [I-D.ietf-idr-tunnel-encaps] to support IPsec tunnels. At IETF 105, IDR and BESS WG held a session to discuss the security issues in these emerging drafts with security directorate people. The security people provided excellent feedback on on how to approach security, privacy, and scaling. The IDR/BESS working members provided provided feedback on the scaling and concepts. As a result of this session, it became evident that each proposal has started with a unique user scenario.

Therefore, this draft simply reviews the technical qualities in terms of: 1) the security and privacy of each technology, and 2) how the technology is managed (manageability) and how the technology scales.

The purpose of this draft to grow our joint understanding of each proposed IPSec tunnel type so that IDR and BESS can make informed decisions. It is non-goal to determine which pf these 3 solutions fits a particular use case using VPN using BGP to pass IPsec tunnel end points.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2. IPSEC deployments

Secure VPNs based on IPsec tunnels began appearing around 2000. IPsec tunnels were used for secure link transport. The IPSEC VPNs utilized IPsec tunnels over physical links or underlay networks to virtual links for these VPNs. These IPsec tunnels can be created by configuring routers with tunnel endpoints and setting up security associations for these tunnels. Automated processes can use IETF network management protocols (NETCONF and RESTCONF) to configure Yang modules in the routers to set up these tunnels.

Enterprise VPNs were created from these secure tunnels. EVPNs and SDWANs have also deployed VPNs using IPSEC.

The BGP client routes which use the tunnel as a pathway also distribute pathway information (endpoint, inner encapsulation, outer encapsulation) via BGP with tunnel attribute [I-D.ietf-idr-tunnel-encaps] For IPsec tunnels, there are three approaches to what security association information is included with the tunnel attribute. (See [I-D.hujun-idr-bgp-ipsec], [I-D.dunbar-idr-sdwan-edge-discovery] and [I-D.sajassi-bess-secure-evpn].

One of the newly defined user scenarios for the secure VPN is the SDWAN. [[I-D.ietf-rtgwg-net2cloud-problem-statement] describes the problems faced by SDWAN. [I-D.ietf-rtgwg-net2cloud-gap-analysis] describes the gaps in IETF technology for this use case. [I-D.dunbar-bess-bgp-sdwan-usage] describes how BGP is used as control plane to setup the SDWAN networks for various SDWAN use cases. SDWAN overlay networks can run over physical forwarding by a wide variety of underlay networks. SDWAN is one of the more recent developments in IPsec based VPNS created by an SDN controller.

The author welcomes additional information on other use cases.

1.3. History of BGP passing Tunnel Endpoints

[RFC5512] defined SAFI to pass tunnel endpoint encapsulation information. However, many operators and vendors preferred to pass this information in a BGP attribute. [I-D.ietf-idr-tunnel-encaps] defines a BGP attribute for tunnels to replace [RFC5512] functionality, but does not address how to use RFC5566 without the encapsulation SAFI. EVPN [RFC8365] also defined tunnel types for encapsulation. The tunnel types registered with IANA (www.IANA.org) list the following tunnel types from [RFC5512], [RFC5566], and [RFC8365]:

- L2TPv3 over IP [RFC5512] [value 1],

- GRE [RFC5512] [value 2]

- Transmit tunnel endpoint [RFC5566][value 3]

- IPsec in tunnel mode [RFC5566] [value 4]

- IP in IP tunnel with IPsec Transport mode [RFC5566][value 5]

- MPLS-in-IP tunnel with IPsec Transport mode [RFC5566][value 6]
- IP in IP [RFC5566] [value 7]
- VXLAN encapsulation [RFC8365][value 8]
- NVGRE encapsulation [RFC8365][value 9]
- MPLS Encapsulation [RFC8365][value 10]
- MPLS in GPE encapsulation [RFC8365] [value 11]
- VXLAN GPE encapsulation [RFC8365] [value 12]

[I-D.ietf-idr-tunnel-encaps] has been created to address deficiencies in RFC5512 [RFC5512]. These deficiencies include: operational costs of using SAFI for tunnel identifiers, inability to specify egress endpoint of tunnel, unclear prefix-tunnel association, and inability to specify inner/outer encapsulation. [I-D.ietf-idr-tunnel-encaps] defines new Sub-TLVs to support inner and outer encapsulation for these encapsulation types, and will become the main reference for these tunnel types.

RFC5566 [RFC5566] defined the IP Tunnel Authenticator Sub-TLV for use in the SAFI, but these recent proposals have suggested different alternatives for replacing the Tunnel Authenticator function.

1.4. Overview of proposals

This section provides a technical overview of the 3 proposals [I-D.hujun-idr-bgp-ipsec], [I-D.dunbar-idr-sdwan-edge-discovery], and [I-D.sajassi-bess-secure-evpn].

[I-D.hujun-idr-bgp-ipsec] proposes 3 new Sub-TLVs: local/remote Prefix Sub-TLV, Public Routing Instance Sub-TLV, and IPSec Configuration Sub-TLV (IPsec-Config). The local/remote prefix Sub-TLV will not be discussed here as it does not clearly align to [I-D.ietf-idr-tunnel-encaps]. The optional Public Routing Instance (PRI) is used instead of a route target community so that local policy can filter routes for a specific community. This feature provides the same feature as a Route target for a pre-configured set of PRIs.

The IPsec Configuration Sub-TLV contains 4 octet opaque value to link the tunnel to the Tunnel Authentication entry found in a security association table on the local node. This table will need to include which tunnel endpoints this security association is valid for. This analysis assumes the IETF protocols NETCONF RESTCONF configure a YANG module that has these security associations.

[I-D.dunbar-idr-sdwan-edge-discovery] proposes UPDATEs from client routers to include the IPsec SA identifiers (ID) to reference the IPsec SA attributes being advertised by separate Underlay Property BGP UPDATE messages. The security association table is built dynamically from the information passed in these Underlay Property BGP Updates plus some local configuration. If a client route can be encrypted by multiple IPsec SAs, then multiple IPsec SA IDs are included in the Tunnel-Encaps Path attribute for the client route. This draft proposes two new Sub-TLVs: IPsec-SA-ID and IPsec-SA-Group. The IPsec-SA-ID is similar to [I-D.hujun-idr-bgp-ipsec] IPSec Config Sub-TLV passing a 2 octet pointer to into a security association table. IPsec-SA-Group Sub-TLV optimizes passing the same information when multiple IPsec SAs with the same inner encapsulation header.

[I-D.dunbar-idr-sdwan-edge-discovery] proposes underlay tunnel topology information for SDWAN in BGP UPDATEs. The information is passed in a combination of NLRI with an SAFI=74 (SDWAN SAFI) and a Tunnel Encapsulation attribute with tunnel type being SDWAN-Underlay.

Security association information for the tunnels in this underlay will be passed in the Tunnel Attribute using in the SDWAN Underlay tunnel type. This new tunnel type which will support the current tunnel Sub-TLV plus the newly proposed IPSec SA Sub-TLV(s). There are two types of IPsec SA Sub-TLVs proposed by [I-D.dunbar-idr-sdwan-edge-discovery], one is for general purpose deployment which requires a full-set of Security Association, including Nonce, Public Key, Proposal and Transform Sub-TLVs in the SDWAN SAFI NLRI (SA-TYPE =2). Another type is for simple deployment which only needs one simple IPsec SA Sub-TLV included (SA-TYPE=1). In addition, it can also include other optional Sub-TLVs like NAT, WAN Port, Geo-location with the SDWAN SAFI route.

[I-D.sajassi-bess-secure-evpn] proposes defines 2 new tunnel types (ESP-Transport and ESP-in-UDP-Transport) and 3 new Sub-TLVS (DIM Sub-TLV, Key-Exchange Sub-TLV, and proposal Sub-TLV) for these new tunnel types. The new Sub-TLVs pass information regarding security associations. The DIM Sub-TLV is required to be supported for the two new tunnel types. As noted above, the SDWAN-WAN-Underlay tunnel type from [I-D.dunbar-idr-sdwan-edge-discovery] supports equivalent features to IPsec-SA, Public-key, and SA-Transforms.

[I-D.dunbar-idr-sdwan-edge-discovery] and [I-D.sajassi-bess-secure-evpn] differ in the information included in the client routes. [I-D.sajassi-bess-secure-evpn] attaches all the IPsec SA information to the actual client routes, whereas the

[I-D.dunbar-idr-sdwan-edge-discovery] only includes the IPsec SA IDs for the client routes. The IPsec SA IDs used by [I-D.dunbar-idr-sdwan-edge-discovery] reference (point) to the SA-Information which is advertised separately. All the SA-Information are attached to routes which describe the SDWAN underlay, such as WAN Ports or Node address.

[I-D.sajassi-bess-secure-evpn] supports tunnel types of ESP-Transport and ESP-in-UDP transport, but not traditional IPsec tunnel types (IPsec in tunnel mode, IP in IP tunnel with IPsec transport, MPLS-in-IP tunnel with transport mode). The use of the new tunnel type could be used in a similar fashion to [I-D.dunbar-idr-sdwan-edge-discovery] to pass SA-information regarding the underlay. [I-D.sajassi-bess-secure-evpn] seems to point to passing client routes upon a rekeying request. This method will increase the amount of BGP traffic passed in the crash or initial start-up in the tunnel encapsulation attribute.

Since [I-D.sajassi-bess-secure-evpn] draft has not recently been updated, it is not clear if the recent changes to [I-D.ietf-idr-tunnel-encaps] are reflected in this draft. [I-D.sajassi-bess-secure-evpn] depends on [I-D.carrel-ipsecme-controller-ike] which received many security comments at IETF 105. Therefore, the author has analyzed [I-D.sajassi-bess-secure-evpn] solutions based on the following assumptions:

- ESP-Transport and ESP-in-UDP would have been aligned with the latest version of the [I-D.ietf-idr-tunnel-encaps],

- Only the DIM Sub-TLV is required to be sent during initialization, PE rekey requests, routing periodic updates, and node restarts (crash/load) for shared security controller policies.

- The multiple policy environments may increase the size of Tunnel Encapsulation attribute as transforms and transform attributes are sent.

1.5. Method of analysis

The things matter to the network operator of IP-SEC VPN in SDWAN: security, manageability, scaling, and privacy. Each deployment of an IPSec VPN may combine different underlay networks with different challenges to security, manageability, scaling and privacy. This analysis compares the basic technologies of these proposals in terms of two groups of features: 1) security and privacy, and 2) manageability and scaling. This analysis drafts looks at each solution based on the strengths are weaknesses of each type.

Analyzing scaling can either be done at the 50,000 foot level or in excruciating detail. This analysis will be at the 50,000 foot level using two example scenarios (small and very large)

Scenario 1: The 3 Route Reflectors (RR) each have 5 client routers per router reflector. The client routers have a potential of 5 tunnels with 1 security association per tunnel. Each client router has 200 routes. The total number of configured tunnels is 20 tunnels per RR cluster and the total number of client routes is 3000. Diagram 1 in section 1.3 showed this simple topology for these route reflectors.

Scenario 2: The 3 Route Reflectors (RR) each have 10,000 client routers. Each client router supports 100 tunnels, 10K routes, and 10 security associations per tunnel. Each Route Reflectors will receive from its client routers a total of 100 million client routes with 1 million tunnels client tunnels (100*10K client routers), and 10 million security associations. The totals for all 3 RR may be up to 3 times this level (300 Million client routes, 3 million tunnels, and 10 million security associations), but it is likely the RR will contain some redundancy. Our scenario focuses on the challenges within a single RR clustr.

The BGP scaling in these two scenarios contrast small IPsec VPNs and very large IPsec VPNs. BGP routing products handle route distribution of over 100 Million routes so this scaling is well within the range of the BGP products.

2. Security and Privacy

During an initial security review of this information, Ben Kaduk made the following comments:

> "First off, when we start to get IPsec configuration via BGP, it’s helpful to think of what other information we get in the same way, and to analyze the effects of misconfiguration or malicious configuration both on IPsec and the broader system. For example, if we are getting NLRI from BGP, then a misconfiguration that gives us parameters that are incompatible with a peer’s is not causing particularly new harm, since we could just as easily be told that peer is unreachable and we wouldn’t try to talk to them anyway. On the other hand, we could be given configuration to use computationally expensive algorithms which would increase the DOS risk in a way that may not (or may!) be already possible. " (email to IDR and BESS WGs after IETF 105)."

The security analysis in this draft assumes that the route distribution for BGP routes is done via a mesh of Route Reflectors which have route reflector clients associated. The IBGP mesh of route reflectors within a domain is assumed to run over secure transport links (such as TLS). If privacy is an issue for BGP route distribution, the TLS encrypts/decrypts the data in the IBGP mesh. If a single AS IBGP mesh of Route Reflectors connects to another AS, the EBGP connection is also over a secure transport (such as TLS).

```
         [Full mesh topology within a IBGP cloud
          LAN used to simply drawing)
                        TLS
           ===============================
          |            |                 |
          RR1---------- RR2 ------------ RR3
    TLS /  |  \      /   |   \         / |  \
       C1  C2  Cn   C1  C2    Cn      C1 C2 Cn
```

Route Reflector Topology

This security topology has the same transport link secure topology as the NETCONF/RESTCONF security of set of NETCONF/RESTCONF servers. The example NETCONF topology is below.

```
           Back-end configuration database
                      TLS
           ===============================
          |            |                 |
        Netconf      Netconf            Netconf
        Client1 -------Client2 -------  Client 3
   TLS  /  |  \      /   |   \         /   |    \
      NS1 NS2 NSn  NS1  NS2  NSn     NS1  NS2   NSn
```

NETCONF Topology

The security aspects of using network management protocols NETCONF or RESTCONF servers to control IPsec SA distribution has been considered as part of SDN-based IPSEC flow consideration (see [RFC8192]). The user data traffic runs over secure IP tunnels whether the configuration is via NETCONF or BGP RR mesh. Figure 3 below shows a Route Reflector topology with BGP sessions in a RR mesh and client traffic over IPSEC tunnels.

Figure 3 - pending [Editor’s note: Large scale topology needs svg drawings]

Figure 4 shows an equivalent topology can be used with NETCONF client-servers. Notice that the NETCONF topology requires a common database behind the network clients to provide the correct configurations. If the NETCONF servers work across administrative domains, a shared database must be developed to provide the appropriate information given the correct policy filters and access (NETCONF NACM).

Diagram 4 - pending [Editor’s note: Large scale BGP topologies need .svg]

There are two parts of the security for control plane traffic: link security and data security. Link security entails making sure the data is secure as the data is transmitted across the link.

Link security in the NETCONF configuration cloud shown in diagram 2 entails making sure the configuration data passes across each of the links. The links from the configuration database (DB in diagram) to each client server must be secured via TLS. The links from the netconf client to the netconf server on the node must be secured via TLS. Data security in the NETCONF configuration cloud entails making sure the data from the configuration data base (DB) travels through the netconf clients (e.g. netconf client1) to the node’s netconf server (e.g. NS1) without change. Data privacy for configuration pathways traffic entails making sure no other party snoops on the data while it travels from configuration database (DB) to the netconf client to the server.

NETCONF client/servers are designed to operate in a single administrative domain. NETCONF client/servers require additional policy filters and checks to run between multiple administrative domains. The Database to NETCONF client link is not standardized by IETF.

Correspondingly, the link security in a BGP RR mesh requires that the data is secure across any link in the BGP RR mesh (RR to/from any RR client or within the RR mesh). Data security for control plane traffic entails ensuring that the data placed into the BGP mesh (from RR clients or RR) arrives at the appropriate destination without change. BGP does not provide data security on control plane traffic as the data may be modified via policy at each node. SBGP does provide data security.

For most networks, physical security of each node and link security is considered sufficient.

The data written into a node using configuration data writes (NETCONF edit-config or RESTCONF PUT/POST) uses the NETCONF client to write to the network server on the client router. The data which is sent from the route-reflector to the RR client routers is sent via BGP, saved in the BGP RIBs, and installed in the router.

The network management protocols (NETCONF/RESTCONF) and BGP both have access policy that controls the data is written into the router. The error handling for incorrect data is different between these two network management protocols (NETCONF/RESTCONF) and BGP. If netconf tries save data with the wrong format, it will provide an error information in the response (rpc-error). The BGP error handling of a malformed Tunnel Attribute in the TLV simply ignores the tunnel attribute while accepting the route.

The common resolution is that NETCONF, RESTCONF, and BGP write error information to a local log. Error reports can be tracked in a Yang module which can be automatically streamed to central controller via an alternate channel NETCONF/RESTCONF logging channel.

[Editorial note: Should I give the details of the NETCONF/RESTCONF logging channel?]

The SDWAN environment or any VPN that uses BGP to transfer tunnel configuration and security association information SHOULD consider augmenting the base BGP Yang model with BGP tunnel encapsulation Yang model for all tunnel types including IPSEC. The logging features or the reporting of the BGP errors can be combined with any error reporting on NETCONF/RESTCONF configuration or any operational state from the tunnel interface. The NETCONF/RESTCONF logging feature providing throttling so any type of error reporting can be configured to be manageable within a large network.

This implies the SDWAN environment should design a BGP Yang model augmenting the BGP base model for the BGP tunnel encapsulation functionality for all tunnel types including IPSEC AND provides logging features the reporting can be the same as NETCONF/RESTCONF.

Given Ben’s rule of thumb, the transmission of the routes, the tunnel end points, and the link to the security association information via the BGP protocol does not cause extra security risks.

The next 3 sections summarize the security and privacy of each technology in terms of:

- what is distributed via netconf,

- what is distributed via BGP,

- link security provided,

- data security provided,

- suggested Yang models that will augment error handling,

- privacy issues.

2.1. Option 1 - Configuration Plus BGP Routes with Tunnel SA IDs

Document: draft-hujun-idr-bgp-ipsec-02.txt [I-D.hujun-idr-bgp-ipsec]

What is distributed via NETCONF: Tunnel Configuration, Security associations, and the mapping of the security association to a tunnel end point (identified y IPsec tunnel identifier), and SA (security association) for the each IPSec tunnel.

What is distributed via BGP: Client Routes with IPsec Sub-TLV per tunnel attribute with IP-SEC tunnel. Optionally, the Public Instance Sub-TLV may augment the BGP tunnel attributes Sub-TLV for tunnel endpoint.

Sub-TLVs added:

IPsec SUB-TLV in IP-Sec Tunnel Attribute (proposed): 4 octet opaque tag.

Public Instance Sub-TLV: identifies the remote instance the Sub-TLV for Tunnel End-Point Identifier takes its address from.

NETCONF Link Security: Distribution is secured by Client-server TLS

Configuration Data security: Configuration clients SHOULD have host and data security. This is beyond NETCONF/RESTCONF security. Client synchronization of data with other clients must have security links and security mechanisms.

Suggested Yang Models for Configuration and Reporting

- Tunnels configuration and operational state

- SA configuration and operational state,

- BGP Tunnel Attribute Yang model augmenting base BGP model with tunnel attributes data and error log. (Tunnel attribute information includes the tunnel attributes plus the mapping of routes to tuples of [tunnel endpoint, security association, and encryption].)

Privacy: Link privacy assumes the ability to encrypt the link data to provide privacy. Node Privacy requires software secure containers within the NETCONF/RESTCONF clients/servers and BGP modules for each of these models.

2.2. Option 2- BGP passes client routes with SA-ID plus NLRI passes underlay SAs

Document: [I-D.dunbar-idr-sdwan-edge-discovery]

What is distributed via NETCONF/RESTCONF or locally configured: Policy and/or templates so that automation may use NLRI with SDWAN SAFI to configure tunnels. This policy may be expressed in as little as 1 line of local configuration.

What is distributed via BGP: Client Routes with tunnel attribute with IPsec Tunnel type, IPSec SA ID(s) which reference Security Association attributes being advertised by SDWAN-Underlay UPDATE. The Sub-TLVs added include:

> IPSec-SA-ID SUB-TLV (proposed): 2 octet pointer into SA table
>
> IPsec-SA-Group SUB-TLV: list of pointers into SA table grouped by inner encapsulation type

The Underlay property is NLRI attached to port addresses or node address with SDWAN SAFI: Includes Site Type, IPSEC-SA-Type, Port-Local-ID, SDWAN-Site-ID, SDWAN Node ID. Depending on the SITE-Type and IPSec-SA-type, this SAFI carries either template references for pre-configured security association (SAs) or full SA information.

> Note: Since the Security association information is carried in a different AFI/SAFI pair, this information may be transmitted in a different BGP update than the client routes with the Tunnel attribute.

Link Security: Distribution is secured between RR to RR clients and between RR in the RR mesh is secured with Transport layer security. If the RR mesh with underlay information is compromised, it does not mean the route with tunnel attributes will be compromised.

Data Security: Data distribution security of tunnel endpoints, SA (security association), routes and mapping (tunnel endpoint, SA, routes) SHOULD have RR and RR Client security on modules processing the data. Full data security (with certificates that the data originated is what arrives at the final destination) for the BGP routes and attributes is beyond the normal mechanism of BGP. These features may be available in SBGP. SBGP signature processing is computationally expensive and requires additional memory space. Synchronization of the routing information on RR (routes, tunnel endpoints, SA-links) and underlay security association information (from AFI/SAFI SDWAN) may be impacted policy that distributes the data.

Technical Note: Many ISPs have chosen to only validate the route origin attribute of the BGP route to insure reduction of "human errors" and some classes of attacks.

Suggested Yang Models for Reporting Errors

- Tunnels configuration and operational state of tunnels,
- SA configuration and operational state of SA information,
- BGP Tunnel Attribute Yang model augmenting base BGP model with tunnel attributes data and error log. (Tunnel attribute information includes the tunnel attributes plus the mapping of routes to tuples of [tunnel endpoint, security association, and encryption].
- Augmentation to base BGP Model to display information passed in NLRI with SDWAN SAFI

Privacy: (Same as option 1)

- Link privacy assumes the ability to encrypt the link data to provide privacy.
- Node Privacy requires secure containers within the netconf/ restconf clients/servers and BGP modules for the BGP control plane data.

2.3. Option 3: Secure EVPN (client routes + SA information)

Document: draft-ietf-sajassi-bess-secure-evpn
[I-D.sajassi-bess-secure-evpn]

What is distributed: Client routes with tunnel attribute with ESP-Transport and ESP-in-UDP-Transport tunnel types.

SUB-TLVs in ESP-Transport and ESP-in-UDP-Transport Attribute (proposed): BASE DIM, Key-Exchange, ESP SA, and Transform Sub-Structure.

This solution would require the Tunnel TLV for the IPsec to contain: Tunnel Endpoint TLV and the DIM TLV.

The DIM SUB-TLV has the following fields:

* ID-length

* Nonce length,

- I-flag

- Flags

- Re-key counter

- Originator ID + (Tenant ID) + (Subnet ID) + (Tenant Address

- Nonce data

Technical Note: The data rate for retransmitting the client routes with the DIM Sub-TLV must be done at the rekeying rate. This automatic re-key counter is distributed with the data.

Link Security: Distribution is secured between RR to RR clients and the RR mesh is secured with transport link security. The regular data distribution of the SA nonce and the rekeying counter provides a potential attack vector for man-in-the middle attacks if the link security is compromised.

Data Security: Data distribution security of tunnel endpoints, SA (security association), routes and mapping (tunnel endpoint, SA, routes) SHOULD have RR and RR Client security on modules processing the data. In addition the processes handling SA information [I-D.carrel-ipsecme-controller-ike] should exist in a protected process.

Full data security for the BGP routes and attributes is beyond the normal mechanism of BGP, but may be available in SBGP. SBGP signature processing is computationally expensive and requires additional memory space. Synchronization of the routing information on RR (routes, tunnel endpoints, SA-links) and underlay security association information (from AFI/SAFI SDWAN) may be impacted policy that distributes the data.

Yang Models for Reporting Errors

o Tunnels configuration and operational state,

o configuration and operational state,

o BGP Tunnel Attribute Yang model augmenting base BGP model with tunnel attributes data and error log. (Tunnel attribute information includes the tunnel attributes plus the mapping of routes to tuples of [tunnel endpoint, security association, and encryption].)

- o Yang models for the operational state in [I-D.carrel-ipsecme-controller-ike].

Privacy: Same as option 1 and 2.

2.4. comparison of security issues

The security of each of these solutions utilizes similar distribution and error reporting. Man in the Middle attacks based on snooping, would need to break the TLS security and encryption for privacy. The [I-D.sajassi-bess-secure-evpn] provides more data directly linked to the routes which could allow an attack vector.

[I-D.hujun-idr-bgp-ipsec] and [I-D.dunbar-idr-sdwan-edge-discovery] provide the route, tunnel information, and link to the SA information. This indirect access to SA information could lessen the attack vector for the tunnel.

[I-D.dunbar-idr-sdwan-edge-discovery] and [I-D.sajassi-bess-secure-evpn] have options send the SA information on unique tunnel types. [I-D.dunbar-idr-sdwan-edge-discovery] placement of the SA data in a NLRI can allow a separate encryption between the SA data and the route/tunnel information.

While all 3 solutions can be used with automated tools (SDN based on simply configuration based), the each solution has benefits and deficits.

3. Manageability and Scaling

Manageability involves how much manual effort is involved to set up IPSec tunnels using each of the three options. The manageability must handle the following: initial set-up of nodes, reporting of status or errors, and rekeying efforts. BGP data distribution and processing of routes to set-up forwarding is stressed during: initial start-up, crash of a RR, and start-up of RR.

The scaling of the system should handle the data distribution and the manageability should handle both the network scenario 1 and network scenario 2. Scenario 1 and scenario 2 both consider one security association per tunnel and 10 security associations per 10. This comparison is given to help understand the impact of rekeying the security associations. [I-D.hujun-idr-bgp-ipsec] would need to send rekeying via NETCONF/RESTCONF, but the rekeying that causes a tunnel to switch security associations can be sent via BGP. [I-D.hujun-idr-bgp-ipsec] use of the NETCONF/RETCONF to send a configuration becomes a bottleneck if the network sizes reach scenario 2. [I-D.dunbar-idr-sdwan-edge-discovery] uses two parallel

BGP NLRI processes where one passes routes and security association identifiers, and the second process sends rekeying based on topology information. Rekeying information is transmitted prior to BGP passes the rekeying of the tunnel. [I-D.sajassi-bess-secure-evpn] passes the rekeying information with the client routes. During initial start-up or RR crash, this rekeying data substantially increases the memory footprint. A continual rekeying process in [I-D.sajassi-bess-secure-evpn] could also cause periodic BGP updates to continue to use bandwidth in the network.

One alternative to the periodic rekeying is to allow the association of more than 1 security association (SA) per tunnel, and allow a local mechanism to switch security associations are a particular time. This analysis looks at the scaling issues of 10 SAs per tunnel allows this analysis to look at the scaling in terms of memory required for this mechanism of rekeying.

The estimate of 10 security associations is admittedly imperfect, but it may help to start the discussion on the memory usage during rekeying.

NETCONF/RESTCONF data distribution scales when the client to netconf/restconf server ratio is low. 1 client per server is best, but 5 servers per client is a low level. 1 client configuring 10K servers on network nodes is beyond most NETCONF/RESTCONF servers. Pushing multiple types of data may also cause stress on the client ability to pull data from the back-end configuration database.

The difference between NETCONF/RESTCONF and BGP mechanisms matter in for SDWAN deployments. NETCONF/RESTCONF is optimized for a single administrative domain and BGP is optimized for inter-domain policy. In SDWAN the nodes are distributed across multiple administrative domains. BGP implementations have many levels of policy. Using BGP each node can be under different RR. Each node can have default SA attributes such as supported encryption algorithms, the nonce, and the public key. The SA ID is only locally significant to the node (or to the port), which is less prone to misconfiguration. BGP also has policy at the route level. Using BGP built-in RT constraint distribution, BGP implementations distribute the SA information to the nodes specified as authorized peers.

3.1. Configuration sizes - used for theoretical comparison

To provide a simple estimate, it is assumed that 100 items needed to be configured in the Yang modules prior to starting the IPsec VPN.

- BGP peer items per node: 20

- o Tunnel configuration items node: 20

- o SA Configuration items per node: 40

- o Monitoring configuration per node: 20

3.2. BGP Route sizes for theoretical comparison

The following estimates are for route and the tunnel Attribute are used for this comparison:

- o average of 4 bytes for IPv4 prefix

- o average of 8 bytes for IPv6 prefix.

3.2.1. Size of Tunnel encapsulation attribute with 1 SA per tunnel endpoint

The space required in the BGP packet for the tunnel attribute per 1 tunnel with 1 Security association (SA) for each of the options is as follows:

- o Tunnel TLV header [4 bytes]

- o Sub-TLV for tunnel endpoint for IPv4 [12 bytes]

- o IP-Sec Sub-TLVs (required) with 1 SA per tunnel endpoint:

  - * Option 1 [I-D.hujun-idr-bgp-ipsec]: 6 octets

  - * Option 2 [I-D.dunbar-idr-sdwan-edge-discovery] 4 octets

  - * Option 3 [I-D.sajassi-bess-secure-evpn] : 35 octets

    - + Sub-TLV header: 3 octets

    - + Dim: 32 octets (header (4), rekey (8), address (8), Nonce (12))

The total space in the tunnel attribute for each type per tunnel endpoint with one security association is the following:

Option 1 [I-D.hujun-idr-bgp-ipsec]: 22 octets

Option 2 [I-D.dunbar-idr-sdwan-edge-discovery]: 20 octets

Option 3 [I-D.sajassi-bess-secure-evpn] : 52 octets

Encapsulation mechanisms such as GRE and VXLAN may add 6-16 octets per tunnel to the Tunnel attribute per tunnel. This addition is due to adding encodings for inner mechanisms (4-12), and outer encodings (2-4)

The total space with encapsulations would then be:

Option 1 [I-D.hujun-idr-bgp-ipsec]: 28-38 octets

Option 2: [I-D.dunbar-idr-sdwan-edge-discovery]:: 26-36 octets

Option 3 [I-D.sajassi-bess-secure-evpn] : 56-68 octets

3.2.2. Size of Tunnel encapsulation attribute with 10 SAs per tunnel

This will be completed in version -01

3.3. Network Scenario 1

This will be completed in version -01

3.4. Network scenario 2

This will be completed in version -01.txt

3.5. Scaling Memory sizes

This section includes scaling for network scenario 1 and 2.

4. Key differences between the options

(to be completed in version -01)

5. Processing of BGP routes

(to be completed in version -01)

6. Future Issues - SBGP and Secure IPSEC VPNs

(to be completed in version -01)

7. Security Considerations

This draft is analysis that includes security and privacy. The draft does not cause any further security issues, but hopes to enhance the security considerations in other drafts.

8. IANA considerations

This draft does not make any requests to IANA for allocations. It is an analysis for review of future allocations in the BGP registry.

9. References

9.1. Normative References

[I-D.carrel-ipsecme-controller-ike]
Carrel, D. and B. Weis, "IPsec Key Exchange using a Controller", draft-carrel-ipsecme-controller-ike-01 (work in progress), March 2019.

[I-D.dunbar-idr-sdwan-edge-discovery]
Dunbar, L., Hares, S., Raszuk, R., and K. Majumdar, "BGP UPDATE for SDWAN Edge Discovery", draft-dunbar-idr-sdwan-edge-discovery-00 (work in progress), July 2020.

[I-D.hujun-idr-bgp-ipsec]
Hu, J., "BGP Provisioned IPsec Tunnel Configuration", draft-hujun-idr-bgp-ipsec-02 (work in progress), March 2020.

[I-D.ietf-idr-tunnel-encaps]
Patel, K., Velde, G., Ramachandra, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", draft-ietf-idr-tunnel-encaps-16 (work in progress), July 2020.

[I-D.sajassi-bess-secure-evpn]
Sajassi, A., Banerjee, A., Thoria, S., Carrel, D., Weis, B., and J. Drake, "Secure EVPN", draft-sajassi-bess-secure-evpn-03 (work in progress), July 2020.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

9.2. Informative References

[I-D.dunbar-bess-bgp-sdwan-usage]
Dunbar, L., Guichard, J., Sajassi, A., Drake, J., Najem, B., and D. Carrel, "BGP Usage for SDWAN Overlay Networks", draft-dunbar-bess-bgp-sdwan-usage-08 (work in progress), July 2020.

[I-D.ietf-rtgwg-net2cloud-gap-analysis]
Dunbar, L., Malis, A., and C. Jacquenet, "Networks Connecting to Hybrid Cloud DCs: Gap Analysis", draft-ietf-rtgwg-net2cloud-gap-analysis-06 (work in progress), May 2020.

[I-D.ietf-rtgwg-net2cloud-problem-statement]
Dunbar, L., Jacquenet, C., and M. Toy, "Dynamic Networks to Hybrid Cloud DCs Problem Statement", draft-ietf-rtgwg-net2cloud-problem-statement-10 (work in progress), May 2020.

[RFC5512] Mohapatra, P. and E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", RFC 5512, DOI 10.17487/RFC5512, April 2009, <https://www.rfc-editor.org/info/rfc5512>.

[RFC5566] Berger, L., White, R., and E. Rosen, "BGP IPsec Tunnel Encapsulation Attribute", RFC 5566, DOI 10.17487/RFC5566, June 2009, <https://www.rfc-editor.org/info/rfc5566>.

[RFC8192] Hares, S., Lopez, D., Zarny, M., Jacquenet, C., Kumar, R., and J. Jeong, "Interface to Network Security Functions (I2NSF): Problem Statement and Use Cases", RFC 8192, DOI 10.17487/RFC8192, July 2017, <https://www.rfc-editor.org/info/rfc8192>.

[RFC8365] Sajassi, A., Ed., Drake, J., Ed., Bitar, N., Shekhar, R., Uttaro, J., and W. Henderickx, "A Network Virtualization Overlay Solution Using Ethernet VPN (EVPN)", RFC 8365, DOI 10.17487/RFC8365, March 2018, <https://www.rfc-editor.org/info/rfc8365>.

Author’s Address

Susan Hares
Hickory Hill Consulting
US

Email: shares@ndzh.com