ROUTING PROTOCOLS FOR ROBUST ENCRYPTED NETWORKS

SAND2007-3055C

Joseph P. Brenkosh
Sandia National Laboratories[1]
Albuquerque, NM 87185
jpbrenk@sandia.gov

ABSTRACT

Most network-centric, geographically dispersed organizations will lease network service from an ISP or long-haul carrier to connect to their various locations. If it is critical that a site maintain the ability to communicate with the rest of the organization, a secondary or backup communications path is a good practice. To enable an automatic switchover between paths, a routing protocol is usually needed. For many commercial and government paths, encryption is required. Many encryption technologies such as IPSec do not directly support routing protocols. This paper examines the issues involved with using routing protocols over IPSec encrypted networks. It then presents solutions for these issues. The solutions presented in this paper are applicable to all those requiring robust encrypted networks, including the warfighter, who may be using a private network. Should the primary network path become unavailable, the secondary path would automatically become the active network connection, allowing the warfighter to focus on the primary mission, not on rerouting network traffic.

INTRODUCTION

Most network-centric, geographically dispersed organizations will lease network service from an ISP or long-haul carrier to connect to their various locations. This is mainly due to the cost of installing and operating the media and gear needed for the network. In most cases the network service leased is based on a service level agreement specifying bandwidth, reliability, etc.

If it is critical that a site maintain the ability to communicate with the rest of the organization, a secondary or backup communications path is a good practice. Depending upon the locations to be connected and the ISP or long-haul carrier, it may be possible to lease network service which has an alternate path that can be used as a backup.

If an alternate path can not be leased directly as a network service, the organization must obtain a backup or alternate path from another ISP or long-haul carrier. The backup path may even be leased from the same firm providing the primary path. The essential requirement is to avoid having any common network gear, fiber, or cable between the two paths which could result in a single point of failure. Because of budget limitations, it is unlikely that the backup path, if purchased separately, will have the same capacity as the primary path. Thus the backup path will only be used when the primary path fails. Switching to the backup path can be accomplished by manually changing routes or having it done automatically by using a routing protocol. The automatic switchover to the backup path is preferred for a number of reasons. The primary reason is that it does not require a knowledgeable person on site 24x7.

To complicate matters, networks carrying data ranging in classification categories from certain types of Unclassified to Top Secret are required to be encrypted in order to insure transmission confidentiality. Although encryption can be performed at any layer of the OSI model, network layer encryption, specifically IP encryption, permits the use of public and private networks in an efficient secure manner.

There are different methods of performing network layer encryption, This paper will focus on methods suitable for an intranet or extranet VPN. An intranet VPN links sites by extending an organization's network across a shared infrastructure. An extranet VPN also links sites, but they are not from the same organization. For these two types of VPNs, all data communication between the sites is encrypted. Peer-to-peer encrypted communication is usually not required. Users rely upon all communications between the sites to be encrypted for them.

IPSec is a method used to provide a full suite of security features for IP data transmission. They include the following:

[1] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

- Data confidentiality – Packets are encrypted before transmitting them across the network.
- Data integrity – An IPSec receiver can authenticate packets sent by an IPSec sender to be certain that they have not been altered during transmission.
- Data origin authentication – An IPSec receiver can authenticate the source of the sent IPSec packets.

IPSec can be implemented in a several ways, depending upon the sensitivity of the data. For non-sensitive US Government data, IPSec can be implemented on the router that an organization uses to connect to its service provider. The remote site or organization must also do the same. Many IPSec implementations on routers allow the administrators to select the encryption algorithm such as DES or AES, and the key distribution method, such as Internet Key Exchange (IKE). Performance is also a consideration, since strong encryption may adversely affect underpowered routers.

Current IPSec implementations support unicast traffic only. Many routing protocols require the use of multicast traffic. This presents a problem for all types of IPSec implementations.

There are two solutions to this problem. The first is to use a Generic Routing Encapsulation (GRE) tunnel to encapsulate all data including routing messages being sent between the IPSec devices. The second method is to use a routing protocol like Border Gateway Protocol (BGP) that can pass routing messages through IPSec.

For all routing protocols discussed, methods for using the preferred path will be presented.

ROUTING THROUGH GRE TUNNELS

This paper presents the GRE tunnel solution with two popular routing protocols, Enhanced Interior Gateway Protocol (EIGRP and Open Shortest Path First (OSPF). As an alternative, floating static routes will also be discussed.

GRE is defined in RFC 2784. It is a tunneling protocol that can support many types of traffic including IP multicast traffic. Thus it provides a means for using routing protocols over networks, such as an IPSec encrypted network, that support only unicast traffic. A GRE packet is an IP packet with the protocol type field set to a value of 47. The data portion of the packet contains a 4 byte GRE header and GRE payload. The GRE payload is the original IP packet.

Routing protocols are configured to use the GRE tunnel to advertise and receive routes and to detect link state changes of the tunnel. By using two tunnels between sites, the backup path can be used, should the preferred path become non-operational. Although the OSPF and EIGRP will perform load balancing, unless the backup path has sufficient bandwidth, load balancing over a link with significantly less bandwidth can cause performance and troubleshooting issues if not properly engineered. The routing protocol should be configured so that the backup path will carry traffic only when the primary path is unavailable.

The following is an example of configuring two GRE tunnels between two sites that are encrypted using IPSec devices.

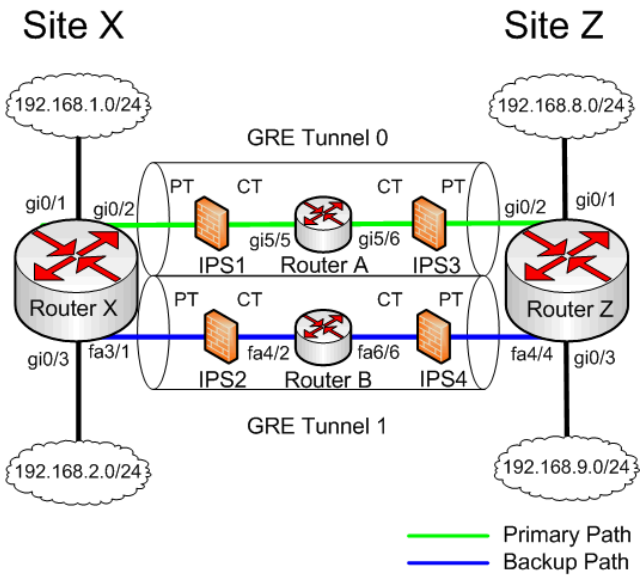

Figure 1. Dual Encrypted GRE Tunnels.

As illustrated in Figure 1. Site X and Site Z are connected by two links, a Gigabit Ethernet link which is the primary path and a Fast Ethernet (100 Mbps) link which is the backup path. Router A is the default gateway for the Gigabit Ethernet link. Router B is the default gateway for the Fast Ethernet link. Each site has a separate IPSec device for performing encryption on both links. They are labeled IPS1 through IPS4. The Plain Text or unencrypted interface of the IPSec device is denoted by *PT*. The Cipher Text or encrypted interface of the IPSec device is denoted by *CT*. Each site supports two user networks. The router interfaces are denoted by the type of interface *gi* for Gigabit Ethernet or *fa* for Fast Ethernet and the Slot/Port Number. Thus *gi0/3* on Router X is Port 3 on a Gigabit Ethernet module inserted in Slot 0. To pass routing protocol information, a separate GRE tunnel is created for the Gigabit Ethernet link and the Fast Ethernet

link. Note this only has to be configured on Router X and Router Z; no other devices need to have any GRE configuration added. Note that this example uses Cisco IOS commands, however, it should be relatively easy to apply this to other routers of other vendors.

OSPF THROUGH A GRE TUNNEL

For OSPF, Router A has the following configuration:

```
interface GigabitEthernet 0/1
description User Network 192.168.1.0
ip address 192.168.1.254  255.255.255.0

interface GigabitEthernet 0/2
description Connection to IPSec Device IPS1
ip address 192.168.4.2 255.255.255.252

interface GigabitEthernet 0/3
description Connection to User Network 192.168.2.0
ip address 192.168.2.254  255.255.255.0

interface FastEthernet 3/1
description Connection to IPSec Device IPS2
ip address 192.168.6.2 255.255.255.252

interface Tunnel 0
ip address 10.1.1.1 255.255.255.252
tunnel source 192.168.4.2
tunnel destination 192.168.5.6

interface Tunnel 1
ip address 10.2.2.1 255.255.255.252
tunnel source 192.168.6.2
tunnel destination 192.168.7.6

router ospf 1
network 192.168.1.0 0.0.0.255 area 0
network 192.168.2.0 0.0.0.255 area 0
network 10.1.1.0 0.0.0.3 area 0
network 10.2.2.0 0.0.0.3 area 0

ip route 192.168.5.0 255.255.255.252 192.168.4.1
ip route 192.168.7.0. 255.255.255.252 192.168.6.1
```

Most of the configuration should be easy to follow. There are some areas to explain however. The interface Tunnel 0 is a virtual interface that has a different address than its physical address. The same is true with Tunnel 1.

For the OSPF routing command section, the number 1 is the process ID for OSPF. This number does not have to match the process ID on the other router. The tunnel interfaces actively participate in OSPF due to the corresponding network 10.1.1.0 and 10.2.2.0 network statements.
Although a GRE tunnel can pass routing information, the tunnel itself needs to be built first. That is the purpose of the two IP routing statements at the end of the configuration.

Router B has a similar configuration:

```
interface GigabitEthernet 0/1
description User Network 192.168.8.0
ip address 192.168.8.254  255.255.255.0

interface GigabitEthernet 0/2
description Connection to IPSec Device IPS3
ip address 192.168.5.6 255.255.255.252

interface GigabitEthernet 0/3
description Connection to User Network 192.168.9.0
ip address 192.168.9.254  255.255.255.0

interface FastEthernet 4/4
description Connection to IPSec Device IPS4
ip address 192.168.6.2 255.255.255.252

interface Tunnel 0
ip address 10.1.1.2 255.255.255.252
tunnel source 192.168.5.6
tunnel destination 192.168.4.2

interface Tunnel 1
ip address 10.2.2.2 255.255.255.252
tunnel source 192.168.7.6
tunnel destination 192.168.6.2

router ospf 1
network 192.168.8.0 0.0.0.255 area 0
network 192.168.9.0 0.0.0.255 area 0
network 10.1.1.0 0.0.0.3 area 0
network 10.2.2.0 0.0.0.3 area 0

ip route 192.168.4.0 255.255.255.252 192.168.5.5
ip route 192.168.6.0. 255.255.255.252 192.168.7.5
```

IPSec device IPS1 has the following configuration:

```
PT Address 192.168.4.1 netmask 255.255.255.252
CT Address 172.16.0.1 netmask 255.255.255.252

Static Route – route 192.168.5.0 to
remote PT 192.168.5.5 - The PT address of remote IPSec device IPS3
remote CT 172.16.0.5 - The CT address of remote IPSec device IPS3
```

CT Gateway 172.16.0.2 – Gi5/5 on Router A. This address will he supplied by the ISP.
IPSec device IPS2 has the following configuration:

PT Address 192.168.6.1 netmask 255.255.255.252
CT Address 172.16.1.1 netmask 255.255.255.252
Static Route – route 192.168.7.0 to
remote PT 192.168.7.5 - The PT address of remote IPSec device IPS4
remote CT 172.16.1.5 - The CT address of remote IPSec device IPS4
CT Gateway 172.16.1.2 – Fa4/2 on Router A. This address will he supplied by the ISP

IPSec device IPS3 has the following configuration:

PT Address 192.168.5.5 netmask 255.255.255.252
CT Address 172.16.0.5 netmask 255.255.255.252
Static Route – route 192.168.4.0 to
remote PT 192.168.4.1 - The PT address of remote IPSec device IPS1
remote CT 172.16.0.1 - The CT address of remote IPSec device IPS1
PT Gateway 172.16.0.6 – Gi5/6 on Router A. This address will he supplied by the ISP

IPSec device IPS4 has the following configuration:

PT Address 192.168.7.5 netmask 255.255.255.252
CT Address 172.16.1.5 netmask 255.255.255.252
Static Route – route 192.168.6.0 to
remote PT 192.168.6.1 - The PT address of remote IPSec device IPS2
remote CT 172.16.1.1 - The CT address of remote IPSec device IPS2
CT Gateway 172.16.1.6 – Gi6/6 on Router A. This address will he supplied by the ISP.

Note that IPSec devices do not need to be configured for the user networks because GRE encapsulates all traffic. Configuration of the IPSec devices is beyond the scope of this paper.

Before a route is entered in the routing table on a router, it must be the best route. The best route is determined by two parameters: administrative distance and a metric that is specific to the routing protocol. For example, the administrative distance for OSPF is 110, and internal EIGRP is 100. The lower the administrative distance, the more believable the routing protocol is considered. Thus if a router had an OSPF and EIGRP route for the same address, the EIGRP route would be the one added to the routing table.

If there are two or more routes for the same address, each having the same administrative distance, then the metric which is specific to the routing protocol is used. For OSPF the metric is cost. Lower cost routes are preferred over higher cost routes. To ensure that the high speed link is used, it may be necessary to influence the value of the metric for a particular interface. For OSPF, the cost of the backup path can be made higher by adding this command under the Tunnel 1 configuration:

ip ospf cost 5000

This should make the cost of Tunnel 1 much higher than Tunnel 0. The value chosen, 5000 in this example, should be higher than the cost of Tunnel 0. Before adding this command, the actual values should be checked with the *show ip route* command or its equivalent.

EIGRP THROUGH A GRE TUNNEL

The configuration for EIGRP is very similar to OSPF. The equivalent EIGRP commands for Router A would be:

router eigrp 25
network 192.168.1.0 0.0.0.255
network 192.168.2.0 0.0.0.255
network 10.1.1.0 0.0.0.3
network 10.2.2.0 0.0.0.3

The commands for Router B would be:

router eigrp 25
network 192.168.8.0 0.0.0.255
network 192.168.9.0 0.0.0.255
network 10.1.1.0 0.0.0.3
network 10.2.2.0 0.0.0.3

The number 25 is an autonomous system (AS) number. For EIGRP they should normally be the same on each router, otherwise route redistribution will need to be performed.

For EIGRP, the metric is calculated by the following formula if default values are used for all EIGRP parameters:
metric = [(10000000 / bw in Kbps) + sum of delays] * 256

Where *bw* is the lowest bandwidth in Kilobits per second of any link being used in the path. The *sum of the delays* is in units of 10 milliseconds. Note that the delay value as displayed by the *show interface* or *show ip eigrp* topology commands is in units of milliseconds. Therefore the value must be divided by ten before using it in the above formula. Thus, to manually influence the preferred route,

add one or more of the following commands under the tunnel interfaces configuration.

bandwidth 1000
delay 200

As can be seen in the above formula, high values for bandwidth lowers the metric and high values for delay increases the metric.

When manually influencing routes, the routers on both ends must be configured similarly or asymmetric routing will occur.

FLOATING STATIC ROUTES THROUGH A GRE TUNNEL

Another option is to not use any routing protocol. This can be accomplished using floating static routes. Floating static routes have not always been a good method of routing multiple WAN links because if there was a problem in anything but the local link, it could not be detected. However, using GRE keepalives solves this problem. GRE keepalives work by embedding a GRE packet in a GRE packet. The embedded GRE packet has the address of the sender. When a GRE keepalive packet is sent, the remote router de-encapsulates the original GRE packet and sends the packet that was embedded back to the sender. When the sender receives its GRE packet back, it knows the link is operational. If it does not, it knows that the link is having problems. To avoid route flapping problems, GRE keepalives can be configured as to the interval between packets sent and also the number of unreturned packets before it declares that the link is not working.

An example of how this can be accomplished is now presented. Using the previous example, the configuration on Router X becomes:

interface Tunnel 0
ip address 10.1.1.1 255.255.255.252
tunnel source 192.168.4.2
tunnel destination 192.168.5.6
keepalive 5 3

interface Tunnel 1
ip address 10.2.2.1 255.255.255.252
tunnel source 192.168.6.2
tunnel destination 192.168.7.6
keepalive 5 3

ip route 192.168.5.0 255.255.255.252 192.168.4.1
ip route 192.168.7.0. 255.255.255.252 192.168.6.1
ip route 192.168.8.0 tu 0
ip route 192.168.9.0 tu 0
ip route 192.168.8.0 tu 1 200
ip route 192.168.9.0 tu 1 200

The keepalive command will cause a GRE keepalive to be sent every 5 seconds, and if it does not receive 3 packets, it declares the link to be non-operational. At that point the static route will be removed from the routing table. The two static routes to 192.168.8.0 will cause Tunnel 0 to be preferred because Tunnel 1 has a higher cost or metric of 200. Note that the rest of the configuration stays the same with the exception of the removal of the OSPF or EIGRP routing commands on the encrypted paths. For this to work properly, Router Z must be similarly configured.

BGP

BGP is the routing protocol used for the internet. It has several important differences from the routing protocols previously discussed. It does not use broadcast or multicast to discover neighbors. Instead BGP uses TCP port 179. The implication of this is that a BGP speaker must have an IP route established before it can contact its peer on TCP port 179. This can be accomplished by using another routing protocol or using static routes.

There are two types of BGP: external BGP (eBGP) and internal BGP (iBGP). External BGP is used when connecting two autonomous systems. An autonoumous system can be a separate organization. It can also be an individual site in the same organization. Autonomous systems have a unique number, with public numbers from 1-64511 and private numbers from 64512-65535. Care must be taken not to use public addresses if they are not owned by your organization. Internal BGP is used by ISPs when they carry other organizations traffic. Although Internal BGP can be used, External BGP is more applicable for connecting two sites in the manner used by this paper.

Using Figure 1 again as an example, only the configuration on Router X and Router Z will change. The changes for Router X are as follows:
Remove all tunnel interfaces
Remove all other routing protocols from the encrypted paths.

Add the following commands:

```
router bgp 65001
network 192.168.1.0
network 192.168.2.0
neighbor 192.168.5.2 remote-as 65002
neighbor 192.168.5.2 ebgp-multihop 255
neighbor 192.168.5.2 route-map PREFFERED-PATH in
```

The first command enables the routing process for autonomous system or AS 65001. The network commands denote the networks that will be advertised. The first neighbor statement defines neighbor 192.168.5.2 to be in AS 65002. The next neighbor command states that neighbor 192.168.5.2 is not directly connected and can be up to 255 (the default) hops away. The last neighbor command instructs BGP to use a route-map called PREFERRED-PATH for incoming BGP routes from neighbor 192.168.5.2

Then the commands for the other interface of the neighbor must be added. Note the absence of the route-map parameter. This is because it is not the primary path.

```
neighbor 192.168.7.2 remote-as 65002
neighbor 192.168.7.2 ebgp-multihop 255
```

Next, an access list is defined for AS 65002 as shown:

```
ip as-path access-list 100 permit ^65002$
```

Finally, the route map is defined. It is called PREFERRED-PATH. Route maps allow routing parameters to be manipulated based upon certain conditions. The route map is

```
route-map PREFERRED-PATH permit 10
match as-path 100
set local-preference 250
```

The first statement names the route map. The permit parameter allows processing to continue as requested by the *set* command if the match criteria succeeds. The number 10 is simply a sequence or line number which could be useful for future editing purposes. The *match* statement determines whether a route it received came from AS 65002, and if it did, the local-preference is increased to 250. For BGP, the default value of local-preference is 100. Setting it higher makes it a more likely candidate to be used. Note that BGP has many tunable parameters for path selection. Local-preference was chosen because it is not vendor specific.

As stated earlier, BGP needs to be able to use TCP. So, the following static routes must remain:

```
ip route 192.168.5.0 255.255.255.252 192.168.4.1
ip route 192.168.7.0. 255.255.255.252 192.168.6.1
```

For Router Z, remove all tunnel references and also all other routing protocols which would be running on the encrypted links. Routing protocols can still be used for the unencrypted internal network. The configuration for Router Z is as follows:

```
router bgp 65002
network 192.168.8.0
network 192.168.9.0
neighbor 192.168.4.2 remote-as 65001
neighbor 192.168.4.2 ebgp-multihop 255
neighbor 192.168.4.2 route-map PREFFERED-PATH in
neighbor 192.168.6.2 remote-as 65002
neighbor 192.168.6.2 ebgp-multihop 255

ip as-path access-list 100 permit ^65001$

route-map PREFERRED-PATH permit 10
match as-path 100
set local-preference 250

ip route 192.168.4.0 255.255.255.252 192.168.5.5
ip route 192.168.6.0. 255.255.255.252 192.168.7.5
```

As stated earlier, it is imperative that both sites agree on the primary and backup paths and configure them appropriately. If the internal network is running a routing protocol such as OSPF, it may be necessary to redistribute the routes learned by BGP into the internal network.

DESIGN CONSIDERATIONS

The design chosen will depend upon many factors. For small networks, where only a few networks need to be reached, static routes and GRE tunnels with keepalives is a simple solution. For large networks, static routes do not scale very well as there could be hundreds of routes. A default route may not always be able to be used when a site connects to several sites via different links.

If both sites are running the same internal routing protocol, GRE tunnels which will allow the routing protocol messages to be passed is a good solution. If performance is an issue, GRE encapsulates every IP packet, not only routing protocol messages. This adds an extra 24 bytes to each packet. Twenty bytes are for the addition IP header and four bytes are for the GRE header.

Also, Path MTU issues may occur if any device has the Do Not Fragment bit set in the IP header.

BGP works well where there the two sites are running different routing protocols. It also does not add any extra overhead. BGP configurations must be carefully scrutinized as a site entering the AS number of a major ISP can cause it to receive traffic destined for that ISP. Also, the routers involved, must be able to support BGP.

CONCLUSION

This paper discusses the issues encountered when attempting to improve the robustness of IPSec encrypted links by using routing protocols. The main issue is that routing protocols are not supported by IPSec devices.

The paper presents three solutions to this problem. They are GRE tunnels using routing protocols and GRE tunnels using floating static routes. BGP is also presented as a solution.

These solutions will allow organizations to maintain a robust encrypted network connection to a remote site. The solutions are applicable to both a network administrator and the warfighter.

BIBLIOGRAPHY

1. Gough, Clare, CCNP BSCI Exam Certification Guide 3rd Edition, Cisco Press, Indianapolis, IN, 2004.

2. Teare, Diane, CCDA Self-Study: Designing for Cisco Internetwork Solutions, Cisco Press, Indianapolis, IN, 2004.