CHAPTER 2

DMVPN Design and Implementation

In designing a VPN deployment for a customer, it is essential to integrate broader design considerations, such as high availability and resiliency, IP multicast, and QoS. This chapter starts with an overview of some general design considerations, followed by sections on implementation, high availability, QoS, and multicast.

Design Considerations

Headend sites are typically connected with DS3, OC3, or even OC12 bandwidth, while branch offices may be connected by fractional T1, T1, E1, T3, or increasingly, broadband DSL or cable access.

To provide redundancy, the branch router should have two or more tunnels to the campus headends. These headend routers can be geographically separated or co-located. For maximum protection, both headend and site redundancy should be implemented. This design guide focuses on the dual DMVPN cloud topology, with both a hub-and-spoke deployment model and a spoke-to-spoke deployment model.

Each deployment model in a dual DMVPN cloud topology has three control planes: the IPsec control plane, the Generic Routing Encapsulation (GRE) control plane, and the routing control plane. Which headend system architecture is chosen determines how each of the control planes is implemented. The following sections provide more detail.

Topology

The following two topologies can be implemented in a DMVPN design:

- Dual hub-dual DMVPN cloud
- Dual hub-single DMVPN cloud

In this design guide, only the dual hub-dual DMVPN cloud topology is discussed. A dual topology allows the network manager slightly easier control over path selection than in a single topology. A dual DMVPN cloud topology can support either a hub-and-spoke deployment model or a spoke-to-spoke deployment model.

The hub-and-spoke deployment model is the most common deployment model. This model is the most scalable, and predominately mimics traditional Layer 2 leased line, Frame Relay, or ATM hub-and-spoke networks. The headend is configured with a multipoint GRE (mGRE) interface, and the branch with a point-to-point (p2p) GRE interface.

The spoke-to-spoke deployment model allows branches to dynamically create tunnels between other branches within the same DMVPN cloud for intercommunication. This deployment model is a fully-meshed topology and requires mGRE interfaces to be configured on both the headend and all branches.

Dual DMVPN Hub-and-Spoke

The hub-and-spoke deployment model in a dual-cloud topology consists of two headend routers, each with one or more mGRE tunnel interface(s) that connect to all branch routers. Each DMVPN cloud represents a unique IP subnet. One DMVPN cloud is considered the primary over which all branch traffic transits. Each branch is configured with p2p GRE tunnel interfaces, with one going to each respective headend. In this deployment model, no tunnels connect one branch to another branch. Traffic between branches passes through the hub router. Routing metrics are used to determine which headend is the preferred path.

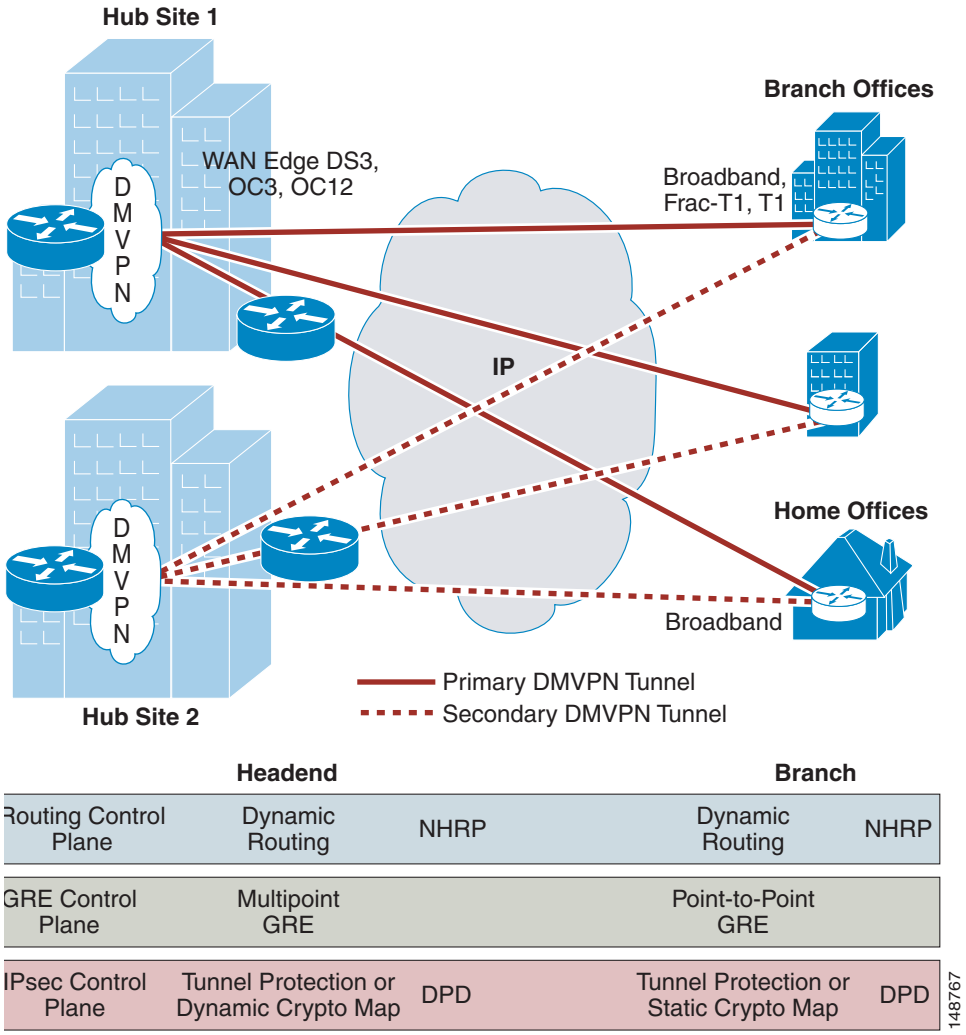
The following two headend system architectures are described in this design guide:

- Single Tier Headend Architecture—Incorporates both the mGRE and crypto functions into a single router processor
- Dual Tier Headend Architecture—Splits the mGRE and crypto functions into two different routers or chasses. This architecture is no longer recommended.

Dual DMVPN Cloud Topology—Hub-and-Spoke Deployment Model (Single Tier Headend Architecture)

Figure 2-1 shows the Single Tier Headend Architecture in a DMVPN deployment.

Figure 2-1 Dual DMVPN Cloud Topology—Hub-and-Spoke Deployment Model (Single Tier Headend Architecture)

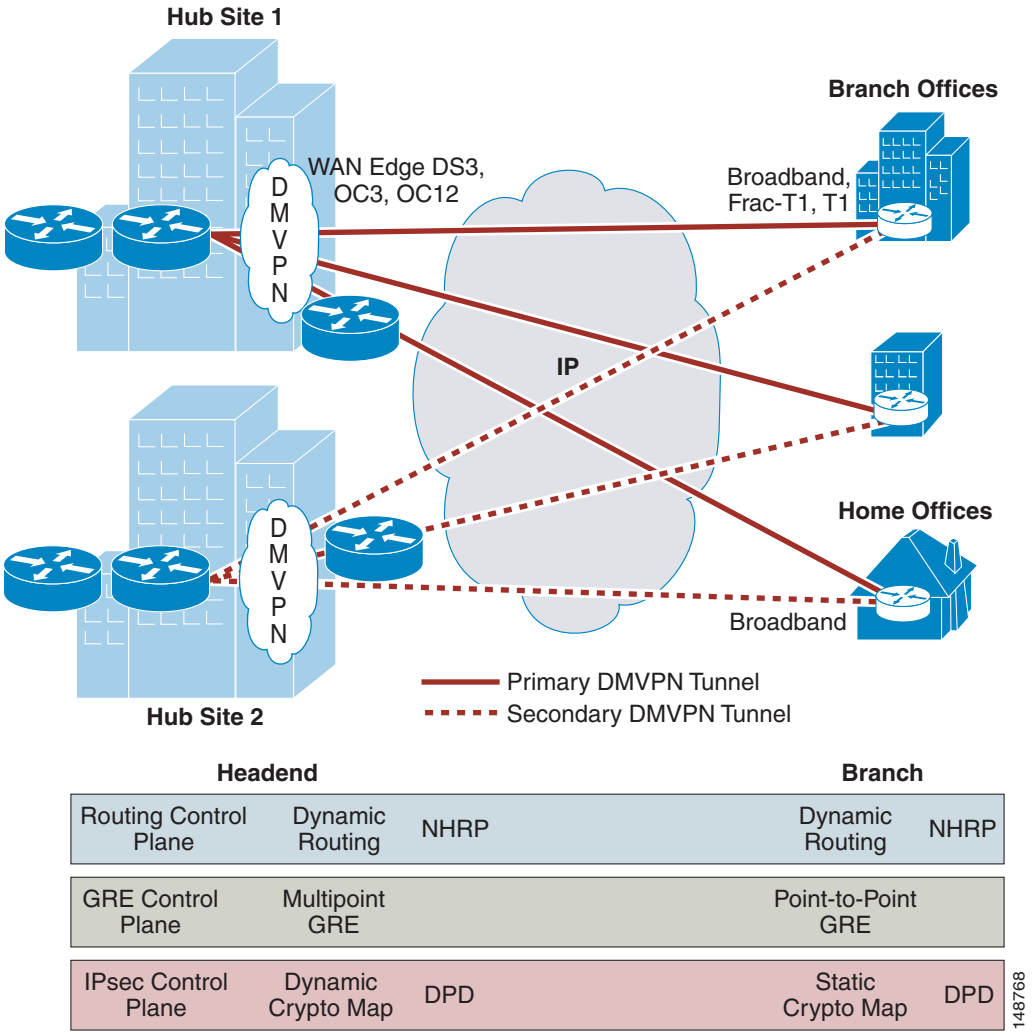


The Single Tier Headend Architecture incorporates all three of the above control planes into a single router. This architecture has an impact on scalability, where the central CPU becomes the gating factor.

Dual DMVPN Cloud Topology—Hub-and-Spoke Deployment Model (Dual Tier Headend Architecture) (this architecture is no longer recommended)

Figure 2-2 shows the Dual Tier Headend Architecture in a DMVPN deployment.

Figure 2-2 Dual DMVPN Cloud Topology—Hub-and-Spoke Deployment Model (Dual Tier Headend Architecture)



The Dual Tier Headend Architecture incorporates the above three control planes into two routers. Both the routing and GRE control planes are housed on one router, while the IPsec control plane is housed on another. Separating the functionality can provide a better scalable solution given various platform limitations; specifically CPU dependencies and resiliency.

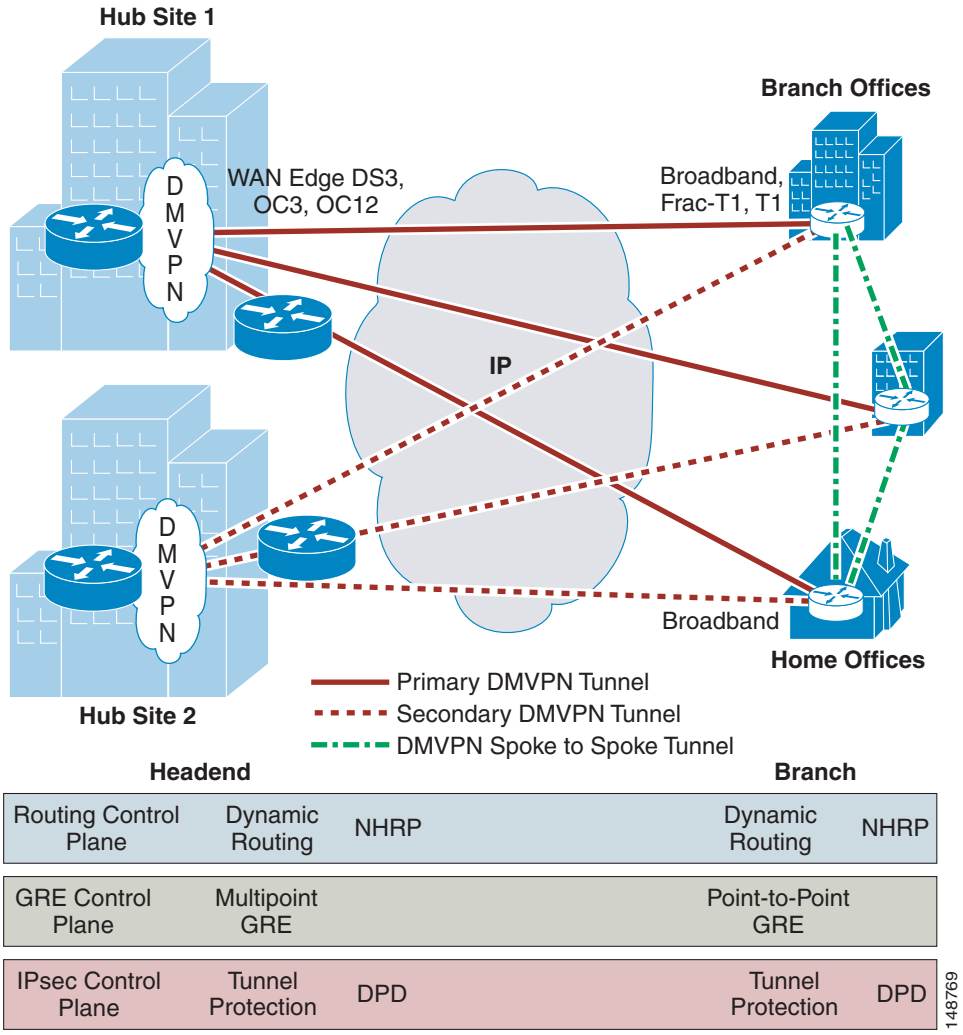
Dual DMVPN Cloud Topology—Spoke-to-Spoke Deployment Model

Spoke-to-spoke deployment in a dual DMVPN topology consists of two headend routers, each with one or more mGRE tunnel interface(s) that connect to all branch routers. Each DMVPN cloud represents a unique IP subnet. One DMVPN cloud is considered the primary, over which all branch traffic transits. On each branch router, there is an mGRE interface into each DMVPN cloud for redundancy. All branch-to-branch communications transit through the primary headend until the dynamic spoke-to-spoke tunnel is created. The dynamic spoke-to-spoke tunnels must be within a single DMVPN cloud or subnet. It is not possible to dynamically create a spoke-to-spoke tunnel between two DMVPN clouds.

Dual DMVPN Cloud Topology—Spoke-to-Spoke Deployment Model (Single Tier Headend Architecture)

Figure 2-3 shows the Single Tier Headend Architecture in a DMVPN deployment.

Figure 2-3 Dual DMVPN Cloud Topology—Spoke-to-Spoke Deployment Model (Single Tier Headend Architecture)



The dual DMVPN cloud topology spoke-to-spoke deployment model with the Single Tier Headend Architecture is very similar to the hub-and-spoke deployment model, with the exception that all GRE interfaces in the headend and the branch are mGRE interfaces. Branch routers can initiate to and accept dynamic tunnels from other branch offices.

IP Addressing

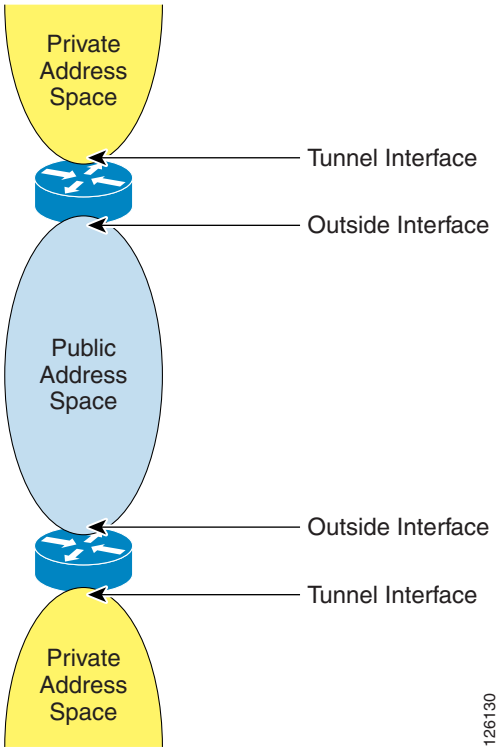
Cisco highly recommends using proper address summarization, which accomplishes the following:

- Conserves router resources, making routing table sizes smaller
- Saves memory in routers and eases troubleshooting tasks

- Simplifies the configuration of routers in IPsec networks

VPNs are used for secure enterprise communications across a shared public infrastructure such as the Internet. Two distinct IP address domains must be considered: the enterprise addressing space, sometimes referred to as the private or inside addresses; and the infrastructure addressing space, also referred to as the service provider, public, or outside addresses. (See Figure 2-4.)

Figure 2-4 Private and Public Address Spaces



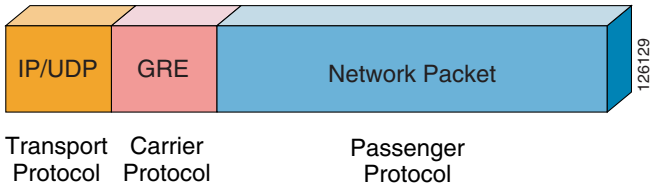
In most DMVPN designs, the outside interface of the router is addressed in the infrastructure (or public) address space, assigned by the service provider. The tunnel interface belongs to the enterprise private network address space. A branch router public IP address is either a statically defined or a dynamically assigned IP address. For a hub-and-spoke deployment model, both the p2p GRE and crypto tunnels are sourced from the public IP address. For a spoke-to-spoke deployment model, the mGRE and crypto tunnels are also sourced from the public IP address. This address is registered with the headend router, which provides a mapping to the branch private address.

Generic Routing Encapsulation—p2p GRE and mGRE Interfaces

Although IPsec provides a secure method for tunneling data across an IP network, it has several limitations. First, IPsec does not support broadcast or IP multicast (IPmc), preventing the use of protocols that rely on these features, such as routing protocols.

Generic Routing Encapsulation (GRE) is a protocol that can be used to “carry” other passenger protocols such as broadcast or multicast IP, as is shown in Figure 2-5.

Figure 2-5 GRE as a Carrier Protocol of IP



Using GRE tunnels in conjunction with IPsec provides the ability to run a dynamic routing protocol or IPmc across the network between the headend(s) and branch offices.

With the p2p GRE over IPsec solution, all traffic between sites is encapsulated in a p2p GRE packet before the encryption process, simplifying the access list used in the crypto map statements. The crypto map statements need only one line permitting GRE (IP Protocol 47). However, in this design, the headend router requires a unique tunnel interface for each branch router, so a large-scale design can have a very large Cisco IOS configuration file on the headend router. For more information on p2p GRE over IPsec designs, see the *Point-to-Point GRE over IPsec Design Guide* at the following URL: http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/P2P_GRE_IPSec/P2P_GRE_IPSec.html.

In DMVPN designs, an mGRE interface is introduced, which serves as a "one-to-many" interface for the creation of multiple hub-and-spoke tunnels that work similarly to a point-to-multipoint Frame Relay interface. Unlike p2p GRE tunnels, the tunnel destination for an mGRE tunnel does not have to be configured. In all DMVPN designs, the headend is configured with an mGRE interface to allow the dynamic creation of tunnels for each branch connected. An mGRE interface does not require a unique tunnel interface, a unique crypto map, or a unique crypto ACL for each branch in the network. mGRE interfaces dramatically reduce the configuration file on each headend router, which is an advantage for large-scale designs when compared to static p2p GRE topologies.

The deployment model chosen determines which type of GRE interface is configured on a branch router. A hub-and-spoke deployment model requires each branch to be configured with a p2p GRE interface. A spoke-to-spoke deployment model requires each branch to be configured with an mGRE interface.

Both p2p GRE and mGRE add to the size of the original data packet, including a four-byte GRE header, a four-byte mGRE tunnel key, and 20 bytes for an additional IP header.

The protocol header for an mGRE packet is four bytes larger than a p2p GRE packet. The additional four bytes constitute a tunnel key value, which is used to differentiate between different mGRE interfaces in the same router. Without a tunnel key, a router can support only one mGRE interface corresponding to one IP network. Tunnel keys allow a branch router to have a different mGRE interface corresponding to each DMVPN cloud in the network topology. A headend router can be configured as well with two mGRE interfaces pointing to each DMVPN cloud for high availability and redundancy.

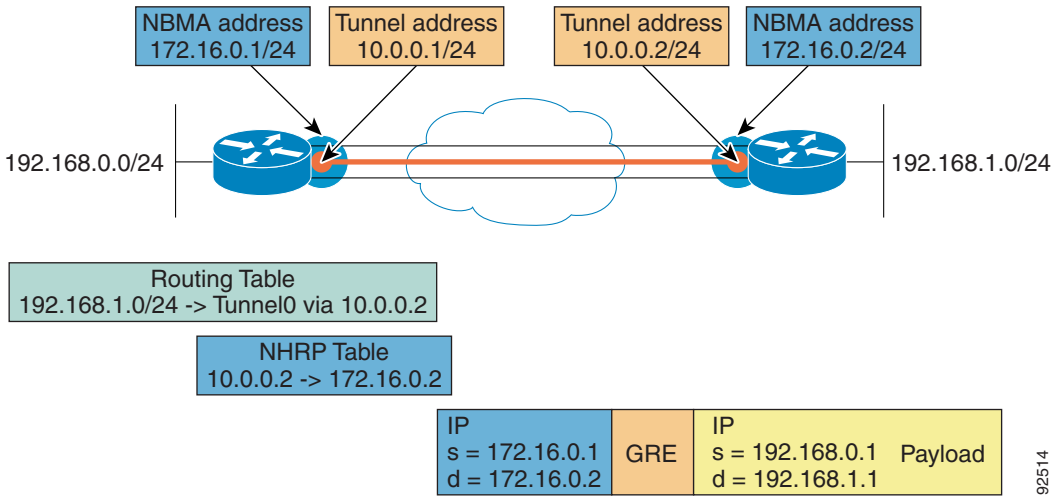
Cisco IOS Software Releases 12.3(13)T, 12.3(11)T3, or later allow multiple mGRE interfaces on a single router to be configured without tunnel keys. Each mGRE interface *must* then reference a unique IP address as its tunnel source.

Next Hop Resolution Protocol

Next Hop Resolution Protocol (NHRP), defined in RFC 2332, is a Layer 2 address resolution protocol and cache, like Address Resolution Protocol (ARP) and Frame Relay Inverse-ARP. NHRP is used by a branch router connected to a non-broadcast, multi-access (NBMA) sub-network to determine the IP address of the "NBMA next hop"; in this case, the headend router or the destination IP address of another branch router.

When a branch router is first established onto a DMVPN network, it registers its IP address with the headend router whose IP address is already pre-configured on the branch router. This registration enables the mGRE interface on the headend router to build a dynamic tunnel back to the registering branch router without having to know the branch tunnel destination through a CLI configuration. NHRP maps a tunnel IP address to an NBMA IP address. NHRP tells the mGRE interface where to tunnel a packet to reach a certain address. When the packet is encapsulated in the mGRE packet, the IP destination address is the NBMA address. Figure 2-6 shows an example of NHRP and mGRE addressing.

Figure 2-6 NHRP and mGRE Addressing



If the destination address is connected to the NBMA sub-network, the headend router is the destination itself. Otherwise, the headend route is the egress router closest to the branch requesting a destination IP address.

Headend and branch routers should be configured with an NHRP holdtime, which sets the length of time that routers instruct other routers to keep their NHRP information. This information is kept in the NHRP cache until the NHRP holdtime expires and the information must be relearned. The default NHRP holdtime is two hours; however, the recommended value is ten minutes. The NHRP cache can be populated with either static or dynamic entries. On the headend router, all entries are added dynamically via registration or resolution requests. The branch router is configured with a static NHRP map pointing to the headend router. To participate in one NHRP registration process, all routers must belong to the same NHRP network by a network ID. The NHRP network ID defines an NHRP domain.

Branch routers must be configured with the NBMA address of the headend router as their next hop server (NHS) to register with the headend router. The branch routers send an NHRP registration through the tunnel to the headend router that contains the tunnel IP address and the NBMA address. The headend router creates an entry in its NHRP cache and returns a registration reply. The branch router now views the headend router as a valid NHS and uses it as a source to locate any other branches and networks in the NHRP domain.

Tunnel Protection Mode

In typical IPsec configurations, dynamic or static crypto maps are configured on the headend and branch routers. These crypto maps specify which IPsec transform set is used and specify a crypto ACL that defines interesting traffic for the crypto map. In Cisco IOS Release 12.2(13)T or later, IPsec profiles are introduced, which share most of the same commands with the crypto map configuration; however, only

a subset of the commands is needed in an IPsec profile. Only commands that pertain to an IPsec policy can be used under an IPsec profile. There is no need to specify the IPsec peer address or the ACL to match the packets that are to be encrypted.

To associate either a p2p GRE or mGRE tunnel with an IPsec profile on the same router, tunnel protection must be configured. Tunnel protection specifies that IPsec encryption is performed after the GRE headers are added to the tunnel packet. With p2p GRE tunnels, the tunnel destination IP address is used as the IPsec peer address. With mGRE tunnels, multiple IPsec peers are possible; the corresponding NHRP-mapped NBMA destination addresses are used as the IPsec peer address. Tunnel protection must be configured on both the headend router and the branch router for a spoke-to-spoke deployment.

If more than one mGRE tunnel is configured on a router that use the same tunnel source address, the **shared** keyword must be added to the **tunnel protection** command on all such tunnel interfaces. Each mGRE tunnel interface still requires a unique tunnel key, NHRP network-ID, and IP subnet address. This is common on a branch router when a dual DMVPN cloud topology is deployed.

Note that GRE tunnel keepalives are not supported in combination with tunnel protection. In addition, tunnel protection cannot be used in a Dual Tier Headend Architecture.

Using a Routing Protocol across the VPN

This design recommends the use of a dynamic routing protocol to propagate routes from the headend to the branch offices. Using a routing protocol has several advantages over the current mechanisms in IPsec Direct Encapsulation alone.

In a VPN, routing protocols provide the same level of benefits as compared to a traditional network, which include the following:

- Propogation of network topology information
- Topology change notification (such as when a link fails)
- Remote peer status

Several routing protocols can be used in a DMVPN design, including EIGRP, BGP, OSPF, RIPv2, and ODR. Designs presented in this design guide use EIGRP as the routing protocol, because EIGRP was used during the scalability testing. EIGRP is recommended as the dynamic routing protocol because of its conservation of router CPU cycles and network bandwidth, as well as its quick convergence times. EIGRP also provides a range of options for address summarization and default route propagation.

Other routing protocols such as OSPF have also been verified, but are not discussed in great detail.Routing protocols increase the CPU utilization on a network device, so this impact must be considered when sizing those devices, especially on the DMVPN headend routers.

Route Propagation Strategy

When a branch connection to the network comes up, the branch router is ready to begin transmitting routing protocol information because it has a static NHRP entry to the headend router. Because the headend router must wait for the NHRP cache to be populated by the branch router, the headend router (NHS) cannot begin sending routing protocol information until after the branch registers its NBMA address.

Crypto Considerations

IPsec supports transport and tunnel encryption modes. Transport mode encrypts only the data portion (payload) of each packet, leaving the source and destination address in the header untouched. The more secure tunnel mode encrypts both the header and payload of the original packet. The difference between these two is that tunnel mode protects the original IP datagram header, and transport mode does not.

Though when IPsec transport mode is used in combination with DMVPN the "IP header that is left in the clear" is the GRE IP header which has the same IP addresses as the IPsec IP header (peers). The original data packet IP header (inside the GRE packet) is fully encrypted, therefore with DMVPN transport mode IPsec is as secure as tunnel mode IPsec. Tunnel mode also adds an additional 20 bytes to the total packet size. Either tunnel or transport mode works in a DMVPN implementation; however, one other restriction with tunnel mode should be understood. If the crypto tunnel transits either a Network Address Translation (NAT) or Port Address Translation (PAT) device, transport mode is required. In addition, this design guide shows configuration examples for implementing DMVPN where the GRE tunnel endpoints are different from the crypto tunnel endpoints (dual Tier), in that case tunnel mode is required. In conclusion, transport mode IPsec is highly recommended and in some cases required with DMVPN.

IKE Call Admission Control

Before Cisco IOS Release 12.3(8)T, there was no means of controlling the number and rate of simultaneous Internet Security Association and Key Management Protocol (ISAKMP) security association (SA) requests received by IKE, which can result in a router being overloaded if more incoming ISAKMP SAs than the processor can handle are initiated. These capabilities are platform-specific. If the processor becomes over-committed, IKE negotiation failures and the constant retransmissions of IKE packets can further degrade router performance.

IKE Call Admission Control (CAC) was introduced in Cisco IOS Release 12.3(8)T to limit the number of IKE authentication of ISAKMP SAs permitted to and from a router. By limiting the amount of dynamic crypto peers that can be created, you can prevent the router from being overwhelmed if it is suddenly inundated with ISKAMP SA requests. The ideal limit depends on the particular platform, the network topology, the application, and traffic patterns. When the specified limit is reached, IKE CAC rejects all new ISAKMP SA requests. If you specify an IKE CAC limit that is less than the current number of active IKE SAs, a warning is displayed, but ISAKMP SAs are not terminated. New ISAKMP SA requests are rejected until the active ISAKMP SA count is below the configured limit.

CAC provides two implementations for limiting IKE SAs that can benefit a DMVPN implementation. First, the normal CAC feature is a global resource monitor that is polled to ensure that all processes including IKE do not overrun router CPU or memory buffers. The user can configure a resource limit, represented by a percentage of system resources from 0 to 100. If the user specifies a resource limit of 90 percent, then IKE CAC drops ISAKMP SA requests when 90 percent of the system resources are being consumed. This feature is valuable on headend routers that can classify and encrypt packets in hardware crypto engines at line rate. It is less useful on branch routers in a hub-and-spoke deployment model, because the branch router typically reaches capacity before being fully loaded with ISAKMP SAs.

The second approach allows the user to configure an IKE authentication limit of ISAKMP SAs (IKE CAC). When this limit is reached, IKE CAC drops all new ISAKMP SA requests. IPsec SA re-key requests are always allowed because the intent is to preserve the integrity of existing sessions. This functionality is primarily targeted at branch routers in a spoke-to-spoke deployment model. By configuring a limit to the amount of dynamic tunnels that can be created to the device, the user can

prevent a router from being overwhelmed if it is suddenly inundated with SA requests. The ideal IKE CAC limit to configure depends heavily on the particular platform and crypto engine (CE), the network topology, and feature set being deployed.

Configuration and Implementation

The configuration issues defined in this chapter are specific to VPN implementation for the dual DMVPN design topology. It is presumed that the reader is reasonably familiar with standard Cisco configuration practices at the command-line interface (CLI) level.

All references to private or public IP addresses correlate to IP Addressing, page 2-5.

For step-by-step instructions, see the following URL:
http://www.cisco.com/en/US/tech/tk583/tk372/tsd_technology_support_protocol_home.html

ISAKMP Policy Configuration

There must be at least one matching ISAKMP policy between two potential crypto peers. The sample configuration below shows a policy using Pre-Shared Keys (PSKs) with 3DES as the encryption algorithm, and SHA as the HMAC. There is a default ISAKMP policy that contains the default values for the encryption algorithm, hash method (HMAC), Diffie-Hellman group, authentication type, and ISAKMP SA lifetime parameters. This is the lowest priority ISAKMP policy.

When using PSK, Cisco recommends that wildcard keys should not be used. However, when implementing a DMVPN design using an IP address obtained dynamically, the use of a wildcard PSK is required. Another approach is the use of Public Key Infrastructure (PKI), also known as Digital Certificates. The example shows two keys configured for two separate crypto peers. The keys should be carefully chosen; “bigsecret” is used only as an example. The use of alphanumeric and special characters as keys is recommended.

The following configuration example shows a static public IP address on the branch router with a static public IP address on the headend router for the crypto peer for either a Single or Dual Tier Headend Architecture:

- Headend router:

```
interface FastEthernet1/0
 ip address 192.168.251.1 255.255.255.0
!
crypto isakmp policy 10
 encr 3des
 authentication pre-share
crypto isakmp key bigsecret address 192.168.161.2
```

- Branch router:

```
interface Serial0/0
  ip address 192.168.161.2 255.255.255.0
!
crypto isakmp policy 10
 encr 3des
 authentication pre-share
crypto isakmp key bigsecret address 192.168.251.1
```

Note the following:

- In a Single Tier Headend Architecture, the configuration above is applied to the headend router.

- I n a Dual Tier Headend Architecture, the configuration above is applied to the crypto headend router.
- In either headend architecture implementing a branch with a dynamic public IP address, a wildcard PSK or PKI must be used on the crypto headend router.

For more information regarding configuring ISAKMP policies, see the following URL: http://www.cisco.com/en/US/docs/ios/12_0/security/command/reference/srike.html.

IPsec Transform and Protocol Configuration

The transform set must match between the two IPsec peers. The transform set names are locally significant only. However, the encryption algorithm, hash method, and the particular protocols used (ESP or AH) must have at least one match. Data compression may also be configured, but it is not recommended on peers with high-speed links. There can be multiple transform sets for use between different peers, with the first match being negotiated. It is important when using multiple transform sets that they be ordered in strongest to weakest order in the **crypto-map** or **crypto profile** configuration.

The following configuration example shows a static public IP address on the branch router, with a static public IP address on the headend router for the crypto peer for either a Single or Dual Tier Headend Architecture:

- Headend router:

```
interface FastEthernet1/0
 ip address 192.168.251.1 255.255.255.0
!
crypto isakmp policy 10
 encr 3des
 authentication pre-share
crypto isakmp key bigsecret address 192.168.161.2
crypto isakmp keepalive 10
!
!
crypto ipsec transform-set vpn-test esp-3des esp-sha-hmac
```

- Branch router:

```
interface Serial0/0
 ip address 192.168.161.2 255.255.255.0
!
crypto isakmp policy 10
 encr 3des
 authentication pre-share
crypto isakmp key bigsecret address 192.168.251.1
crypto isakmp keepalive 10
!
!
crypto ipsec transform-set vpn-test esp-3des esp-sha-hmac
```

Note the following:

- In a Single Tier Headend Architecture, the configuration above is applied to the headend router.
- In a Dual Tier Headend Architecture, the configuration above is applied to the crypto headend router.
- In either headend architecture implementing a branch with a dynamic public IP address, a wildcard PSK or PKI must be used.

For more information on transform sets and configuring crypto maps, see the following URL: http://www.cisco.com/en/US/docs/ios/12_2/security/command/reference/srfipsec.html.

Tunnel Protection Configuration

Tunnel protection can be used when the GRE tunnel and the crypto tunnel share the same endpoints. Because of this restriction, tunnel protection is applicable only to the Single Tier Headend Architecture.

In early versions of IPsec configurations, dynamic or static crypto maps specify which IPsec transform set (encryption strength and Diffie-Hellman group) and also specify a crypto access list, which defines interesting traffic for the crypto map. As of Cisco IOS Software Release 12.2(13)T, the concept of an IPsec profile exists. The IPsec profile shares most of the same commands with the crypto map configuration, but only a subset of the commands is needed in an IPsec profile. These commands pertain to an IPsec policy that can be issued under an IPsec profile; there is no need to specify the IPsec peer address or the ACL to match the packets that are to be encrypted. The IPsec peer address and the crypto match ACL are provided automatically by the tunnel protection code.

A sample IPsec profile is shown in the following example:

- Headend router:

```
crypto ipsec transform-set ESE esp-3des esp-sha-hmac
  mode transport
!
crypto ipsec profile VPN-DMVPN
  set transform-set ESE
!
```

- Branch router:

```
crypto ipsec transform-set ESE esp-3des esp-sha-hmac
  mode transport
!
crypto ipsec profile VPN-DMVPN
  set transform-set ESE
!
```

The IPsec profile is associated with a tunnel interface using the **tunnel protection ipsec profile** *profile-name* command, also first introduced in Cisco IOS Software Release 12.2(13)T. The **tunnel protection** command can be used with mGRE and p2p GRE tunnels. With p2p GRE tunnels, the tunnel destination address is used as the IPsec peer address. With mGRE tunnels, multiple IPsec peers are possible; the corresponding NHRP-mapped NBMA destination addresses are used as the IPsec peer addresses. Crypto access lists that define the interesting traffic no longer need to be configured.

If more than one mGRE tunnel is configured on a router (for example, on a branch router with dual DMVPN clouds), it is possible to reference the same tunnel source address on each tunnel interface. In this case, the **shared** keyword is used in the **tunnel protection** command on both interfaces. This does not mean that the two mGRE tunnels are hosting the same DMVPN cloud; each tunnel interface still requires a unique tunnel key, NHRP network-ID and IP subnet.

Dynamic Crypto Map Configuration

The dynamic crypto map is required only in a dual tier architecture where tunnel protection cannot be used. The following configuration examples show a dynamic public IP address on the branch router with a static public IP address on the headend router using a Dual Tier Headend Architecture:

- Headend router:

```
interface FastEthernet1/0
 ip address 192.168.251.1 255.255.255.0
!
crypto isakmp key bigsecret address 0.0.0.0 0.0.0.0
```

```
!
crypto dynamic-map dmap 10
 set transform-set vpn-test
!
!
crypto map dynamic-map local-address FastEthernet1/0
crypto map dynamic-map 10 ipsec-isakmp dynamic dmap
```

- Branch router:

```
interface Serial0/0
 ip address dhcp
!
crypto isakmp key bigsecret address 192.168.251.1
!
crypto map static-map local-address Serial0/0
crypto map static-map 20 ipsec-isakmp
 set peer 192.168.251.1
 set transform-set vpn-test
 match address vpn-static2
 set security-association level per-host

ip access-list extended vpn-static2
 permit gre any host 192.168.251.1
```

Note the following:

- On the branch router the crypto ACL had to match **any** as the tunnel source address, because there the Serial0/0 IP address cannot be known at configuration time. This will cause problems on the headend since the reverse crypto ACL, which the headend would use, will match **any** as the IP destination which would match all branch routers. This can be mitigated with using the **set security-association level per-hos**t command under the crypto map on the spoke. This will cause the spoke to dynamically limit the crypto ACL sent to the hub to source host (current IP address of Serial1/0 interface). It is easier to just use **ipsec profile** and **tunnel protection** as described in other sections.
- On the headend router, a dynamic crypto map is used with a wildcard PSK to allow a crypto peer with the public dynamically-served IP address of the branch router.
- In a Dual Tier Headend Architecture, the configuration above is applied to the crypto headend router.

For a more complete description of the various crypto configuration commands, see the *Cisco IOS Security Command Reference*.

Applying Crypto Maps

Crypto maps are required only when a Dual Tier Headend Architecture is used. The crypto map is applied on the routers outside the public interface. The branch router must also be configured with a static crypto map when a Dual Tier Headend Architecture is used because the encryption tunnel destination differs from the GRE tunnel destination.

The following configuration example shows a public dynamic IP address on the branch router with a static public IP address on the headend router for the crypto peers for a Dual Tier Headend Architecture:

- Headend router:

```
interface FastEthernet1/0
  ip address 192.168.251.1 255.255.255.0
  crypto map dynamic-map
!
```

- Branch router:

```
interface Serial0/0
 ip address dhcp
 crypto map static-map
!
```

In a Dual Tier Headend Architecture, the configuration above is applied to the crypto headend router.

mGRE Configuration

The configuration of mGRE allows a tunnel to have multiple destinations. The configuration of mGRE on one side of a tunnel does not have any relation to the tunnel properties that might exist at the exit points. This means that an mGRE tunnel on the hub may connect to a p2p tunnel on the branch. Conversely, a p2p GRE tunnel may connect to an mGRE tunnel. The distinguishing feature between an mGRE interface and a p2p GRE interface is the tunnel destination. An mGRE interface does not have a configured destination. Instead the GRE tunnel is configured with the command **tunnel mode gre multipoint**. This command is used instead of the **tunnel destination** x.x.x.x found with p2p GRE tunnels. Besides allowing for multiple destinations, an mGRE tunnel requires NHRP to resolve the tunnel endpoints. Note, tunnel interfaces by default are point-to-point (p-p) using GRE encapsulation, effectively they have the **tunnel mode gre** command, which is not seen in the configuration because it is the default.

The mGRE configuration is as follows:

```
!
interface Tunnel0
 bandwidth 1536
 ip address 10.62.1.10 255.255.255.0
 tunnel source Serial0/0
 tunnel mode gre multipoint
!
```

Tunnel Interface Configuration—Hub-and-Spoke Only

This section illustrates the tunnel interface configurations using a branch static public IP address.

The following configuration example shows a static public IP address on the branch router with a static public IP address on the headend router for either a Single or Dual Tier Headend Architecture:

- Headend router:

```
interface FastEthernet1/0
 ip address 192.168.251.1 255.255.255.0
!
interface Tunnel0
 bandwidth 1536
 ip address 10.62.1.1 255.255.255.0
 tunnel source FastEthernet1/0
 tunnel mode gre multipoint
!
```

- Branch router:

```
interface Serial0/0
 ip address 192.168.161.2 255.255.255.252
```

```
!
interface Tunnel0
 bandwidth 1536
 ip address 10.62.1.194 255.255.255.0
 tunnel source Serial0/0
 tunnel destination 192.168.251.1
!
```

Note that this configuration applies only in a Single Tier Headend Architecture.

Tunnel Interface Configuration—Dynamic Spoke-to-Spoke

This section illustrates the tunnel interface configurations using a branch dynamic public IP address.

The following configuration example shows a dynamic public IP address on the branch router with a static public IP address on the headend router for the mGRE tunnel for a Single Tier Headend Architecture:

- Headend router:

```
interface FastEthernet1/0
 ip address 192.168.251.1 255.255.255.0
!
interface Tunnel0
 bandwidth 1536
 ip address 10.62.1.1 255.255.255.0
 tunnel source FastEthernet1/0
 tunnel mode gre multipoint
!
```

- Branch router:

```
interface Serial0/0
 ip address dhcp
!
interface Tunnel0
 bandwidth 1536
 ip address 10.62.1.10 255.255.255.0
 tunnel source Serial0/0
 tunnel mode gre multipoint
!
```

Note the following:

- In a Single Tier Headend Architecture, the configuration above is applied to the headend router.
- In a Dual Tier Headend Architecture, the configuration above is applied to the mGRE headend router. The mGRE headend router has a different static public IP address than the crypto headend router. The mGRE headend router sends all outbound mGRE traffic to the branch through the crypto headend.

NHRP Configuration

NHRP provides a mapping between the inside and outside address of a tunnel endpoint. These mappings can be static or dynamic. In a dynamic scenario, a next-hop server (NHS) is used to maintain a list of possible tunnel endpoints. Each endpoint using the NHS registers its own public and private mapping with the NHS. The local mapping of the NHS must always be static. It is important to note that the branch points to the inside or protected address of the NHS server.

The NHRP hold time is also used to determine how often a branch sends an NHRP registration to its configured headends (NHS). The spoke will send NHRP registrations every 1/3 of its configured NHRP hold time. This default NHRP registration timer setting can be changed independently of the NHRP hold time with the **ip nhrp registration timeout** ***value*** command, where ***value*** is in seconds and should not be set lower than 75 seconds. This automatic sending of NHRP registrations is used to keep the branch's NHRP registered mapping up to date on the headend (NHS).

Although spoke-to-spoke voice (VoIP) over DMVPN is not generally recommended because of QoS concerns, the NHRP hold time should be longer than the duration of the majority of calls. The hold timer should not be so long that spoke-to-spoke sessions are idle on average. This recommendation is especially true for low-end routers where the software imposes a lower limit on the number of crypto tunnels. An overall balance between idle tunnels and excessive re-caching can be achieved by setting the idle time to 600 seconds. Note, NHRP checks to see if the spoke-spoke tunnel is still being used during the last 120 seconds before a dynamic (spoke-spoke) mapping would expire. If it is still being used then NHRP will proactively send an NHRP resolution request soliciting an NHRP resolution reply to refresh the dynamic spoke-spoke NHRP mapping before it expires, otherwise it will allow the NHRP mapping to expire and the spoke-spoke tunnel to be removed. The configurations are as follows:

- Headend router:

```
!
interface Tunnel0
 description NHRP with mGRE
 ip address 10.0.0.1 255.255.255.0
 ip mtu 1400
 ip nhrp map multicast dynamic
 ip nhrp network-id 12345
 ip nhrp holdtime 600
 tunnel source FastEthernet0/0
 tunnel mode gre multipoint
 tunnel key 100000
 tunnel protection ipsec profile DMVPN shared
!
```

- Branch router:

```
!
interface Tunnel0
 description NHRP with p2p GRE
 ip address 10.0.0.11 255.255.255.0
 ip mtu 1400
 ip nhrp map 10.0.0.1 192.168.251.1
 ip nhrp network-id 12345
 ip nhrp holdtime 600
 ip nhrp nhs 10.0.0.1
 tunnel source FastEthernet0/0
 tunnel destination 192.168.251.1
 tunnel key 100000
 tunnel protection ipsec profile DMVPN shared
!
```

Note In the spoke tunnel configuration **ip nhrp map multicast 192.168.251.1** is not required, because this is a p-pGRE tunnel and any IP multicast packets forwarded out the tunnel interface will automatically be GRE encapsulated and sent to the single tunnel destination (192.168.251.1). On multipoint GRE (mGRE) tunnels it is required to have a static NHRP multicast mapping for each NHS unicast mapping so that IP multicast packets forwarded out the tunnel interface will be encapsulated to the one or more configured headend (NHS).

Routing Protocol Configuration

Because the DMVPN cloud is a non-broadcast, multi-access network, some considerations must be made when running dynamic routing protocols. This is particularly true when implementing a spoke-to-spoke design. Many routing protocols have an IP multicast mechanism that is used to discover other participating nodes. Static multicast maps are configured on branch routers pointing to the public address of the hub. The hub router is configured with a dynamic multicast map. This allows the hub and spokes to exchange broadcast information, but does not permit spokes to hear the broadcasts from other spokes.

EIGRP Configuration

EIGRP is the preferred routing protocol when running a DMVPN network. The deployment is straightforward in a pure hub-and-spoke deployment. The address space should be summarized as much as possible, and in a dual cloud topology, the spokes should be put into an EIGRP stub network. As with all EIGRP networks, the number of neighbors should be limited to ensure the hub router can re-establish communications after a major outage. If the DMVPN subnet is configured with a /24 network prefix, the neighbor count is limited to 254, which is a safe operational limit. Beyond this number, a compromise is required to balance re-convergence with recovery. In very large EIGRP networks, it may be necessary to adjust the EIGRP hold time to allow the hub more time to recover without thrashing. However, the convergence time of the network is delayed. This method has been used in the lab to establish 400 neighbors.

Spoke-to-spoke DMVPN networks present a unique challenge because the spokes cannot directly exchange routing information with one another, even though they are on the same logical subnet. This limitation requires that the headend router advertise subnets from other spokes on the same subnet. This would normally be prevented by split horizon. In addition, the advertised route must contain the original next hop as learned by the hub router. A new command (**no ip next-hop-self**) was added to allow this type of operation. The **no ip split-horizon eigrp AS** and **no ip next-hop-self eigrp AS** commands are only configured on DMVPN headend routers, configuring these commands on DMVPN branch routers can cause EIGRP to become unstable. The following configurations detail a typical EIGRP configuration. Note that the outside address space of the tunnel should not be included in any protocol running inside the tunnel.

- Headend router:

```
!
interface Tunnel0
 description Tunnel0
 bandwidth 100000
 ip address 10.56.0.1 255.255.252.0
 no ip redirects
 ip hold-time eigrp 1 35
 no ip next-hop-self eigrp 1
 ip nhrp authentication test
 ip nhrp map multicast dynamic
 ip nhrp network-id 105600
 ip nhrp holdtime 600
 no ip split-horizon eigrp 1
tunnel source GigabitEthernet0/0/0
 tunnel mode gre multipoint
 tunnel key 10000
!
router eigrp 1
 network 10.0.0.0
 no auto-summary
!
```

- Branch router:

```
!
interface Tunnel0
 description Tunnel0
 bandwidth 100000
 ip address 10.56.0.3 255.255.252.0
 no ip redirects
 ip hold-time eigrp 1 35
 ip nhrp authentication test
 ip nhrp map 10.56.0.1 192.168.201.1
 ip nhrp map multicast 192.168.201.1
 ip nhrp network-id 105600
 ip nhrp holdtime 600
 ip nhrp nhs 10.56.0.1
 ip nhrp registration timeout 120
 tunnel source FastEthernet0/0/0
 tunnel mode gre multipoint
 tunnel key 10000
!
router eigrp 1
 network 10.0.0.0
 no auto-summary
 eigrp stub connected
!
```

OSPF Configuration

Configuring OSPF over a DMVPN network has some of the same limitations as OSPF over other types of networks. Historically, a single OSPF area should not contain more than 50 routers, and there should not be more than three areas on a router. Although current routers have stronger processors, the additional overhead of encryption and NHRP negates much of this. For this reason, the 50 router limit per area should be observed. In addition, because only the hub is in direct communications with all of the branches, it must be configured as the designated router (DR) on the DMVPN subnet. There is not typically a backup designated router (BDR). A BDR is possible if a second hub is placed on the same subnet. This is common in a single-cloud, dual-hub topology.

The mGRE tunnel on the hub router must be configured as an OSPF broadcast network to allow the selection of a DR. Each spoke router is configured with an OSPF priority of 0 to prevent a spoke from becoming the DR. In addition, if the spoke is configured with p2p GRE and the hub is mGRE, the hello timer on the spoke should be changed from the default of 10 seconds to 30 seconds to match the hello timers on the mGRE interface. The tunnel IP MTU must match on all GRE interfaces that are OSPF-adjacent. In addition, OSPF areas running over DMVPN should be stubby or totally stubby areas to reduce LSA flooding over the WAN. The configuration is as follows:

- Headend router:

```
!
interface Tunnel0
 description dmvpn tunn
 ip address 10.173.20.1 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication secret
 ip nhrp map multicast dynamic
 ip nhrp network-id 10203
 ip nhrp holdtime 600
```

```
 ip ospf network broadcast
 ip ospf priority 200
 tunnel source GigabitEthernet0/1.201
 tunnel mode gre multipoint
 tunnel key 123
 tunnel protection ipsec profile dmvpn
!
router ospf 10
 network 10.173.20.0 0.0.0.255 area 10
 area 10 stub no-summary
!
```

- Branch router:

```
!
interface Tunnel0
 ip address 10.173.20.21 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication secret
 ip nhrp map multicast 192.168.201.1
 ip nhrp map 10.173.20.1 192.168.201.1
 ip nhrp network-id 10203
 ip nhrp holdtime 600
 ip nhrp nhs 10.173.20.1
 ip route-cache flow
 ip ospf network broadcast
 ip ospf priority 0
 load-interval 30
 tunnel source GigabitEthernet0/0.201
 tunnel mode gre multipoint
 tunnel key 123
 tunnel protection ipsec profile dmvpn
!
router ospf 10
 network 10.173.20.0 0.0.0.255 area 10
 area 10 stub no-summary
!
```

In hub-and-spoke only networks, it is possible to reduce the OSPF load by using a point-multipoint network type on the headend router and point-point network type on the branch routers. In this case, there is no need to elect a DR router on the DMVPN subnet. The headend router serves as the master for the subnet. The branches consider the headend as the only path off the subnet, thus simplifying the Dijkstra algorithm for the OPSF area.

RIPv2 Configuration

RIPv2 over DMVPN is possible. Configurations are not shown. If RIPv2 is used for the routing protocol, the **no ip split-horizon** command must be configured on the hub mGRE tunnel interface if spoke-to-spoke traffic is to be permitted, even via the hub. By default, RIPv2 uses the original IP next hop instead of itself when advertising routes out the same interface from where it learned them; therefore, there is no need for a “next-hop-self” configuration. When spoke-to-spoke tunnels are in use, auto-summary must be disabled. The configuration is as follows:

- Headend router:

```
!
interface Tunnel0
 description dmvpn tunn
 ip address 10.173.20.1 255.255.255.0
```

```
 no ip redirects
 ip mtu 1400
 no ip split-horizon
 ip nhrp authentication secret
 ip nhrp map multicast dynamic
 ip nhrp network-id 10203
 ip nhrp holdtime 600
 ip ospf network point-to-multipoint
 ip ospf hello-interval 30
 tunnel source GigabitEthernet0/1.201
 tunnel mode gre multipoint
 tunnel key 123
 tunnel protection ipsec profile dmvpn
!
router rip
 version 2
 network 10.0.0.0
 no auto-summary
!
```

- Branch router:

```
!
!
interface Tunnel0
 ip address 10.173.20.21 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication secret
 ip nhrp map multicast 192.168.201.1
 ip nhrp map 10.173.20.1 192.168.201.1
 ip nhrp network-id 10203
 ip nhrp holdtime 600
 ip nhrp nhs 10.173.20.1
 ip ospf network point-to-point
 ip ospf hello-interval 30
 tunnel source GigabitEthernet0/0.201
 tunnel mode gre multipoint
 tunnel key 123
 tunnel protection ipsec profile dmvpn
!
router rip
 version 2
 network 10.0.0.0
 no auto-summary
```

High Availability

High availability (HA) provides network resilience and availability in the event of a failure. This section provides some designs for highly-available DMVPN implementations. High availability is covered in much more detail in the *IPsec VPN Redundancy and Load Sharing Design Guide* at the following URL: http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/VPNLoad/VPN_Load.html.

Common Elements in all HA Headend Designs

To provide a level of resiliency in the VPN design, Cisco recommends that at least two tunnels and two headends be configured on each branch router. Regardless of DMVPN topology or deployment model, each branch router should have a tunnel to a primary headend and an alternate tunnel to a secondary headend router.

Under normal operating conditions, both the primary and secondary tunnels are up and have routing protocol neighbors established. The routing protocol maintains both paths. The routing protocol can be configured to make the secondary tunnel (headend) as a less preferred path.

A common concern in all HA headend resilient designs is the number of routing protocol neighbors. Many redundant neighbor relationships increase the time required for routing convergence.

Routing protocol convergence is a common element in all HA headend designs. However, each deployment model has unique methods of achieving HA via routing protocol convergence.

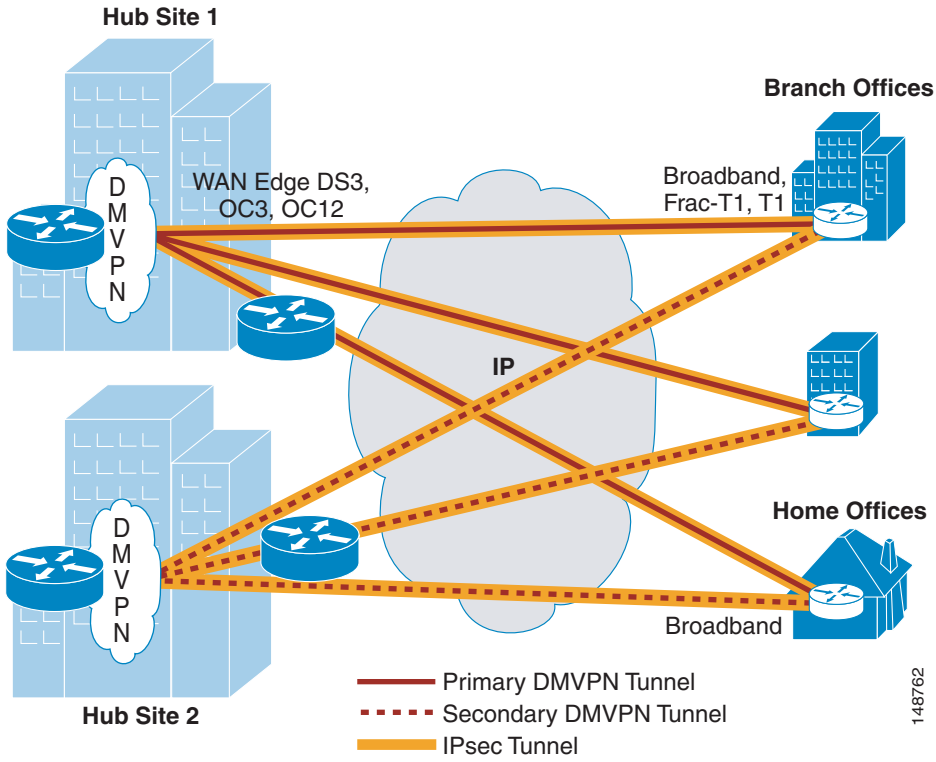
Dual DMVPN Cloud Topology—Hub-and-Spoke Deployment Model

This chapter describes two headend system architectures for a dual DMVPN cloud topology. Each headend architecture described handles HA uniquely. The following sections describe HA in a hub-and-spoke deployment model with various headend architectures.

Hub-and-Spoke Deployment Model—Single Tier Headend Architecture

Figure 2-7 shows a hub-and-spoke deployment model with the Single Tier Headend Architecture for a typical HA scenario.

Figure 2-7 Hub-and-Spoke Deployment Model—Single Tier Headend Architecture



If a failure occurs at one of the headend devices, the routing protocol detects that the route through the primary tunnel is no longer valid and, after convergence, the route through the secondary tunnel is used. When the primary tunnel is available again, traffic is routed back to the primary tunnel, because it is the preferred route in the routing metrics. The headend resiliency design presented here allows for failure of a single headend router, with proper failover to surviving headend routers, regardless of IP subnet or DMVPN cloud.

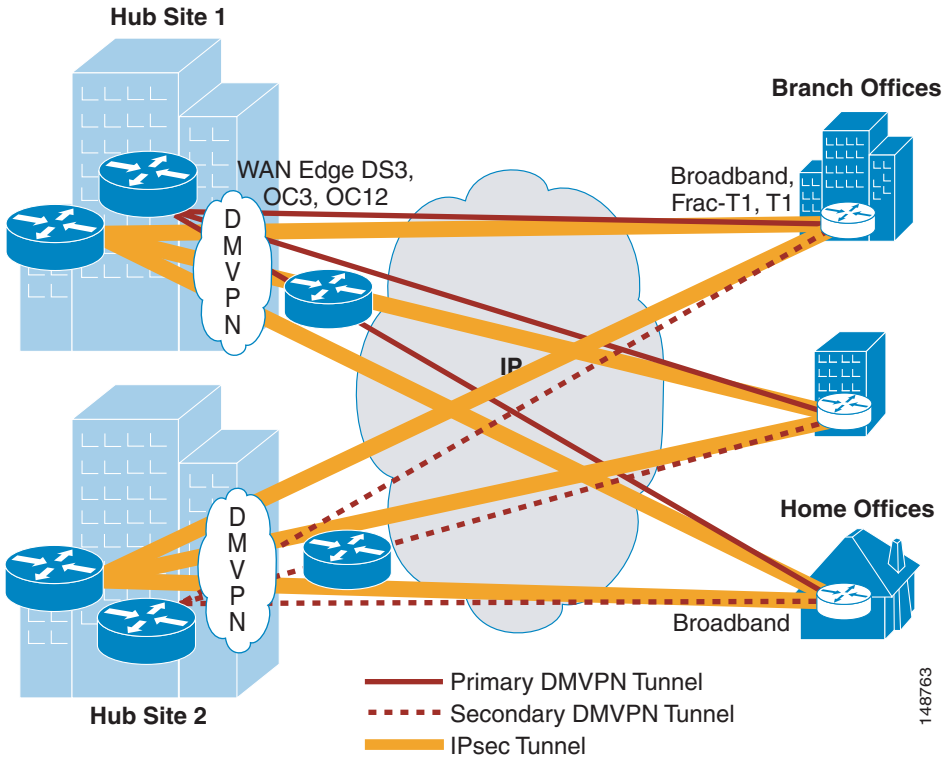
It is possible to configure more than one mGRE interface on a hub router. Two mGRE tunnels can be configured with the same tunnel source if the **shared** keyword is specified on the **tunnel protection** command. However, the two mGRE tunnels would still create two separate DMVPN networks using a unique tunnel key, NHRP network-ID, and IP subnet on each tunnel interface.

The typical branch router has two or more tunnel interfaces to two or more VPN headends. All tunnels from the branch to the headend routers are up. The routing protocol determines which tunnel is passing user traffic. The various paths in this design are configured with slightly different metrics to provide preference between the tunnels. The routing metric should be consistent both upstream and downstream to prevent asymmetric routing.

Hub-and-Spoke Deployment Model—Dual Tier Headend Architecture

Figure 2-8 shows a hub-and-spoke deployment model with the Dual Tier Headend Architecture for a typical HA scenario.

Figure 2-8 Hub-and-Spoke Deployment Model—Dual Tier Headend Architecture



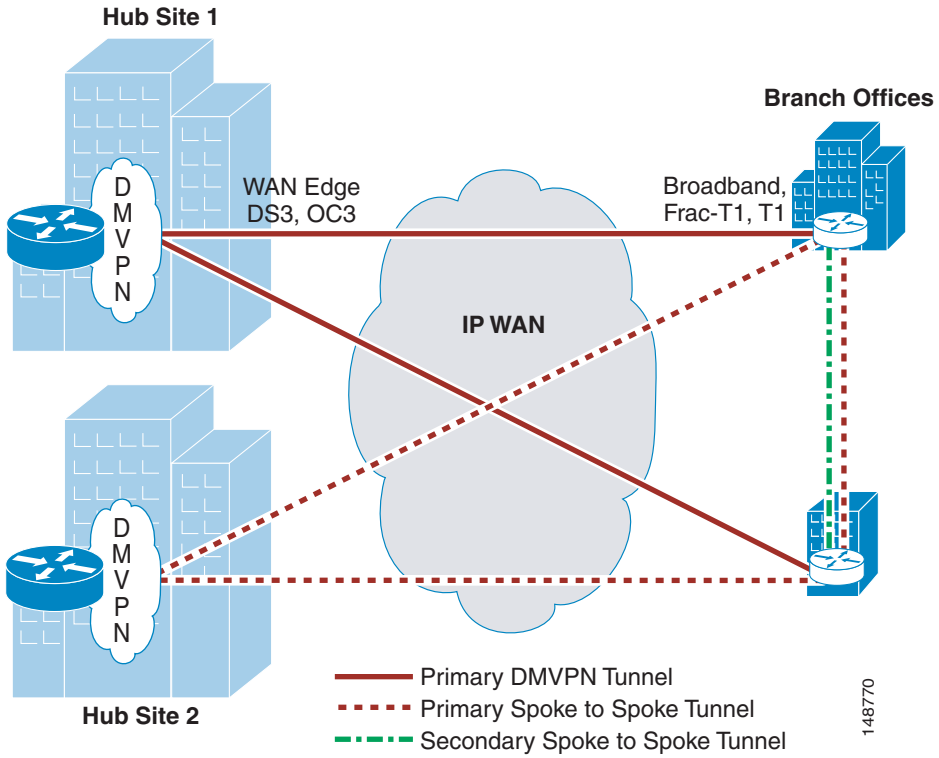
If Dual Tier Headend Architecture is implemented, the crypto functionality is separated from the GRE and routing protocol functions. This provides additional paths in the event of a failure providing there is connectivity between both pairs of headends. Conversely, this architecture also contains additional devices that could possibly fail. The dual tier architecture does not allow crypto profiles or tunnel protection to be implemented. The crypto configurations on the branch require manual mapping to both possible crypto headends. Failover configurations to allow either crypto headend to be used are discussed in the IPsec Direct Encapsulation Design Guide at the following URL: http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/Dir_Encap.html.

A failure of the GRE tunnel is handled in the same manner as the Single Tier Headend Architecture. In this situation, a dynamic routing protocol chooses the backup DMVPN subnet.

Dual DMVPN Cloud Topology—Spoke-to-Spoke Deployment Model

Figure 2-9 shows a spoke-to-spoke deployment model in a typical HA scenario.

Figure 2-9 Spoke-to-Spoke Deployment Model



In addition to routing protocols determining the primary and secondary path similar to the hub-and-spoke deployment model, the spoke-to-spoke deployment model includes other HA considerations.

For a spoke-to-spoke tunnel to be dynamically created between branch routers, each branch must traverse through a single DMVPN cloud to obtain both the primary route and the proper NHRP address for the other branch. In other words, a spoke-to-spoke tunnel between Branch 1 and Branch 2 in the Figure 2-9 must be connected to a single DMVPN cloud. For a spoke-to-spoke tunnel to be created from Branch 1 to Branch 2 through DMVPN Cloud 1, Branch 1 and Branch 2 have a static NHRP map to Headend 1 to obtain the IP addresses of each branch dynamically, and a route in the routing table of each branch so that they can communicate.

If Headend 1 fails, routing protocol convergence occurs just as in the hub-and-spoke deployment model. Branch 1 is now routed through Headend 2 to reach Branch 2. Headend 2 is in a separate DMVPN cloud, which means a new spoke-to-spoke tunnel between Branch 1 and Branch 2, now through DMVPN Cloud 2, must be created. The original spoke-to-spoke tunnel pointing through DMVPN Cloud 1 remains as long as the NHRP hold time and IPsec SA timers are active. When the NHRP hold time and IPsec SA timers expire, the original spoke-to-spoke tunnel terminates. When Headend 1 recovers, routing converges back. Traffic is placed back on the original DMVPN subnet, and the IPsec SAs used for the spoke-to-spoke session on the backup DMVPN subnet are torn down after the timers have expired.

QoS

To support latency-sensitive traffic applications, it may be necessary to configure QoS. QoS and IPsec have been integrated as part of the Cisco Voice and Video Enabled IPsec VPN (V3PN) technology. For more information, see the *Voice and Video Enabled IPsec VPN (V3PN) Design Guide* at the following

URL: http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/V3PN_SRND/V3PN_SRND.html.

Ideally, a service provider has implemented a QoS configuration on both the link to the head-end campus location as well as on the access link to each branch router. There are Cisco Powered Network Service providers offering this QoS capability for transporting encrypted voice and data. They can be located by using the Cisco Powered Network—Find Recommended Service Providers utility at the following URL: http://www.cisco.com/pcgi-bin/cpn/cpn_pub_bassrch.pl. Search with the criteria of "IP VPN-Multiservice" in the dialog box.

However, the enterprise customer may be in the position where QoS must be provisioned for the campus head-end to branch router. The following section outlines how this can be accomplished.

QoS in a Hub-and-Spoke Deployment Model

In a hub-and spoke deployment, branch routers can implement a QoS service policy on the outside physical interface and obtain congestion feedback from the physical interface where the interface clock rate is the actual uplink rate. Examples of this are serial T1 or E1 interfaces or Frame Relay. Branch routers that are attached by way of broadband or Ethernet hand-off from a service provider customer premises equipment (CPE) route need to implement Hierarchical Class-Based Weighted Fair Queueing (HCBWFQ) on the outside physical interface, and queue within a shaped rate that is derived from the contracted uplink rate.

From the headend perspective, all the spoke routers are often accessed from a high-speed interface; for example, a Gigabit Ethernet or OC3 link that presents the possibility of the hub router overrunning the access link of the spoke router. In a Frame Relay network, the solution to this problem is implementing Frame Relay traffic shaping on a hub router.

In a DMVPN network, the solution is the Dynamic Multipoint VPN Hub Support by Quality of Service Class feature. This feature is available beginning in Cisco IOS Software Release 12.4 (9)T and later for Cisco 7200 Series routers and 7301 routers, and provides the ability to implement a per tunnel (per security association/per branch) QoS service policy.

Note Amore extensive QoS feature, Per-tunnel QoS over DMVPN for hub to spoke traffic was implemented in 12.4(20)T and later for 7200, 7301, ISR, and ASR 1000 class of routers.

An example of how this configuration is implemented is demonstrated as follows. This example assumes that there are two spoke routers with inside LAN network addresses of 10.0.92.0/24 and 10.0.94.0/24, as shown.

- Spoke router 1

```
!
hostname vpn-jk2-831-1
!
interface Ethernet0
 description Inside LAN Interface
 ip address 10.0.92.1 255.255.255.0
!
end
```

- Spoke router 2

```
!
hostname vpn-jk2-831-2
!
interface Ethernet0
```

```
 description Inside LAN Interface
 ip address 10.0.94.1 255.255.255.0
!
end
```

In the head-end configuration to support these branches, each branch has an IP access list configured to match packets destined from any network at the campus to the remote inside LAN network at the respective branch.

```
!
hostname HEAD_END
!
ip access-list extended B000
 permit ip any 10.0.92.0 0.0.0.255
ip access-list extended B001
 permit ip any 10.0.94.0 0.0.0.255
```

Also, a class map is configured for each branch to reference the appropriate IP extended access list:

```
class-map match-all B001-class
 match access-group name B001
class-map match-all B000-class
 match access-group name B000
!
```

Class maps are defined to identify the traffic to prioritize to each branch:

```
!
class-map match-all VOICE
 match ip dscp ef
class-map match-any CALL-SETUP
 match ip dscp af31
 match ip dscp cs3
class-map match-any INTERNETWORK-CONTROL
 match ip dscp cs6
!
```

These class maps are referenced by the policy map configuration. The VOICE class is configured for nine G.729 voice calls, assuming the downlink to each branch is a Frame Relay or HDLC-encapsulated T1 link from the service provider to the branch router. This is the “child” service policy in an HCBWFQ configuration.

```
policy-map branch-policy
 class CALL-SETUP
  bandwidth percent 2
 class INTERNETWORK-CONTROL
  bandwidth percent 5
 class VOICE
  priority 504
 class class-default
  fair-queue
  random-detect
!
```

The “parent” service policy in an HCBWFQ configuration is defined. It is assumed that each branch is connected to the service provider at a T1 (1.54 Mbps) data rate. The shaper is configured at a percentage of that rate, in this case 85 percent, to accommodate some degree of jitter in the arrival rate of all packets because of queuing and buffering within the service provider network. The 85 percent example is a conservative value, and can likely be incremented to avoid wasting bandwidth. However, the goal is to never present more packets than the link can handle. Do not configure a shaper value that allows the service provider to drop packets indiscriminately. It is assumed that the service provider has not applied

any QoS to the access link. If the service provider is using a Cisco router and has the default configuration of "fair-queue" on the T1 link, the shaped rate may exceed 90–95 percent because Weighted Fair Queue is precedence-aware by default and thus is inherently QoS-aware.

```
policy-map Shaper-1544K-all
 description 1544K * .85 = 131K
 class B000-class
  shape average 1310000 13100
  service-policy branch-policy
 class B001-class
  shape average 1310000 13100
  service-policy branch-policy
!
! ... and so on for all branches
!
```

On the headend hub router, the mGRE tunnel interface is configured for *qos pre-classify* because the service policy is matching on the destination IP address in the original unencrypted IP header. The service policy, however, is applied to the outside interface, and the packets are encrypted when the QoS matching decision is invoked. Configuring *qos pre-classify* gives the service policy the ability to match on the clear text values.

```
!
interface Tunnel0
...
 qos pre-classify
!
interface FastEthernet0/1.100
 description Outside interface
...
 service-policy output Shaper-1544K-all
```

Performance characteristics for this configuration are provided in Chapter 4, "Scalability Test Results (Unicast Only)."

The following router output is from a **show policy-map** displayed during the performance scale testing.

```
show policy-map interface
 GigabitEthernet0/1

  Service-policy output: Shaper-1544K-all

    Class-map: b000-class (match-all)
      158299 packets, 48139994 bytes
      30 second offered rate 1299000 bps, drop rate 0 bps
      Match: access-group name b000
      Traffic Shaping
           Target/Average   Byte   Sustain   Excess    Interval  Increment
             Rate           Limit  bits/int  bits/int  (ms)      (bytes)
          1310000/1310000   3275   13100     13100     10        1637

        Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
        Active Depth                         Delayed   Delayed   Active
        -      12        158304    48140792  147759    45034738  yes

      Service-policy : branch-policy

        Class-map: CALL-SETUP (match-any)
          0 packets, 0 bytes
          30 second offered rate 0 bps, drop rate 0 bps
          Match: ip dscp af31 (26)
            0 packets, 0 bytes
            30 second rate 0 bps
```

```
    Match: ip dscp cs3 (24)
      0 packets, 0 bytes
      30 second rate 0 bps
    Queueing
      Output Queue: Conversation 73
      Bandwidth 2 (%)
      Bandwidth 26 (kbps) Max Threshold 64 (packets)
      (pkts matched/bytes matched) 0/0
  (depth/total drops/no-buffer drops) 0/0/0

  Class-map: INTERNETWORK-CONTROL (match-any)
    5 packets, 870 bytes
    30 second offered rate 0 bps, drop rate 0 bps
    Match: ip dscp cs6 (48)
      5 packets, 870 bytes
      30 second rate 0 bps
    Match: access-group name IKE
      0 packets, 0 bytes
      30 second rate 0 bps
    Queueing
      Output Queue: Conversation 74
      Bandwidth 5 (%)
      Bandwidth 65 (kbps) Max Threshold 64 (packets)
      (pkts matched/bytes matched) 5/870
  (depth/total drops/no-buffer drops) 0/0/0
    QoS Set
      dscp cs6
        Packets marked 5

  Class-map: VOICE (match-all)
    118416 packets, 18709728 bytes
    30 second offered rate 502000 bps, drop rate 0 bps
    Match: ip dscp ef (46)
    Queueing
      Strict Priority
      Output Queue: Conversation 72
      Bandwidth 50 (%)
      Bandwidth 655 (kbps) Burst 16375 (Bytes)
      (pkts matched/bytes matched) 110494/17458052
      (total drops/bytes drops) 0/0

  Class-map: class-default (match-any)
    39878 packets, 29429396 bytes
    30 second offered rate 795000 bps, drop rate 0 bps
    Match: any
    Queueing
      Flow Based Fair Queueing
      Maximum Number of Hashed Queues 64
  (total queued/total drops/no-buffer drops) 6/0/0
       exponential weight: 9

class    Transmitted      Random drop      Tail drop    Minimum Maximum  Mark
         pkts/bytes       pkts/bytes       pkts/bytes    thresh  thresh  prob
    0   39344/29075840        0/0              0/0          20      40  1/10
    1       0/0               0/0              0/0          22      40  1/10
    2     536/353824          0/0              0/0          24      40  1/10
    3       0/0               0/0              0/0          26      40  1/10
    4       0/0               0/0              0/0          28      40  1/10
    5       0/0               0/0              0/0          30      40  1/10
    6       0/0               0/0              0/0          32      40  1/10
    7       0/0               0/0              0/0          34      40  1/10
 rsvp       0/0               0/0              0/0          36      40  1/10
```

```
Class-map: b001-class (match-all)
  158223 packets, 48128074 bytes
  30 second offered rate 1301000 bps, drop rate 0 bps
  Match: access-group name b001
  Traffic Shaping
       Target/Average   Byte   Sustain   Excess    Interval  Increment
         Rate           Limit  bits/int  bits/int  (ms)      (bytes)
      1310000/1310000   3275   13100     13100     10        1637

    Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
    Active Depth                       Delayed   Delayed   Active
    -      8         158231    48130666  148135    45116434  yes

. . . and so on. There is one instance of a shaper and class map for each branch.
However the display is terminated for brevity.
```

Note the following:

- The offered rate of 1299000 bps is approaching the shaper rate of 1310000; the downlink to this branch is fully utilized
- The shaper is engaged in the b000-class, and there are currently 12 packets (queue-depth) queued
- Packets are matched in the VOICE class, the CALL-SETUP has no matches as the test traffic profile does not contain any packets marked CS3/AF31
- The INTERNETWORK-CONTROL has matches, and these packets are EIGRP hello packets
- The default class has matches in both IP Precedence 0 and 2. The traffic profile has both DSCP BE (Best Effort) and DSCP AF21 (IP precedence 2); WRED is enabled and these markings are displayed in the appropriate counters.

QoS in a Spoke-to-Spoke Deployment Model

In a spoke-to-spoke deployment model, branch routers can also be configured with the Dynamic Multipoint VPN Hub Support by Quality of Service Class feature. Cisco IOS Release 12.4 (9)T or later is required for spokes consisting of Cisco 800, 1700, 1800, 2600XM, 2800, 3700, 3800 Series routers. For Cisco 831 Series spokes, Cisco IOS release 12.3 (11) T10 is required.

The challenge in the spoke-to-spoke model is that traffic is originating from the hub site as well as possibly one or more spoke sites. If the hub site and the spoke sites all configure their QoS shaper at the access link data rate (or some percentage of that rate, as described in the previous section), there is the possibility that the collection of the hub and spokes send sufficient traffic to overrun a single spoke. Determining the appropriate values for both the shapers as well as the priority and bandwidth queues is a challenge at best.

For this reason, a QoS-enabled service provider offering QoS on the access link to the branch or spoke routers is the best possible solution.

IP Multicast

IPmc has two areas of concern. The first is the ability to scale the solution such that a large number of listeners may join a stream. The second is the restrictions in functionality required as a result of the point-to-multipoint GRE interface.

Scalability testing with IPmc and IPsec encryption indicates that there are issues with packet loss because of the instant replication of many packets. IPmc replication generates new headers for each of the NHRP mapped destinations. The payload is not changed. Each packet is referenced as a pointer that

links the header and the payload to downstream software process, such as encryption. After a single packet is replicated, the list of pointers that is passed to encryption can overrun the inbound RxRing on the crypto card. This is a certainty if the number of joined destinations exceeds the size of the RxRing. It is a strong possibility if the stream has a high pps rate or multiple multicast streams are flowing, because the encryption process may not be able to drain the RxRing before receiving the burst of packets.

For example, consider a design using the Cisco Catalyst 6500 with VPN Shared Port Adaptor (SPA), and configuring 1000 p2p GRE over IPsec tunnels to branch offices. If each branch office is joined to a single multicast stream, the VPN SPA must replicate each multicast packet 1000 times, one per VPN tunnel. Assuming the Sup720 can sustain the replication speed of the stream, many packets (up to 1000) arrive at the input queue of the VPN SPA, causing overruns or dropped packets. If the customer has IPmc requirements, see the *Multicast over IPsec VPN Design Guide* for appropriate scalable designs at the following URL:
http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/V3PNIPmc.html.

Because the DMVPN network is a non-broadcast subnet, special situations must be considered before deploying IPmc over a DMVPN network. First, spokes that are members of the same subnet are not able to form PIM adjacencies with one another. In a spoke-to-spoke topology, the RPF checks do not allow multicast flows to transit through the hub. This prevents IPmc between spokes because flows always need to traverse the hub. In addition, and for the same reasons, a spoke should not be used as a rendezvous point. If a multicast rendezvous point (RP) is placed at the hub site then IP multicast can be configured so that the RP will hair-pin the IP multicast packets coming in from a spoke site and forward them back out the mGRE tunnel interface on the headend to other spokes that are requesting the IP multicast stream.

IPmc over DMVPN works in a hub-and-spoke deployment, providing the number of joined branches does not exceed the RxRing limit of the encryption engine. If an IPmc stream originates from a branch location, an RP must be deployed at the hub site in order for other spoke site clients to receive the stream. The following configuration examples show multicast over DMVPN with the IP Multicast source at the hub site.

- Hub router:

```
!
interface Tunnel0
 description dmvpn tunnel
 ip address 10.173.20.1 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip pim sparse-mode
 ip pim nbma-mode
 ip nhrp authentication secret
 ip nhrp network-id 10203
 ip nhrp map multicast dynamic
 ip nhrp holdtime 600
tunnel source GigabitEthernet0/1.201
 tunnel mode gre multipoint
 tunnel key 123
 tunnel protection ipsec profile dmvpn
!
```

- Spoke router:

```
!
interface Tunnel0
 description dmvpn tunnel
 ip address 10.173.20.10 255.255.255.0
 no ip redirects
 ip mtu 1400
```

```
 ip pim sparse-mode
 ip pim nbma-mode
 ip nhrp authentication secret
 ip nhrp map 10.173.20.1 192.168.201.1
 ip nhrp map multicast 192.168.201.1
 ip nhrp network-id 10203
 ip nhrp holdtime 600
 ip nhrp nhs 10.173.20.1
 load-interval 30
 tunnel source Serial0/0
 tunnel mode gre multipoint
 tunnel key 123
 tunnel protection ipsec profile dmvpn
!
```

Interactions with Other Networking Functions

Other networking functions such as NAT, PAT, DHCP, and firewall considerations apply to designing a DMVPN network. This section describes these functions.

Network Address Translation and Port Address Translation

Although NAT and PAT can result in an added layer of security and address conservation, they both present challenges to the implementation of an IPsec VPN. ISAKMP relies on an individual IP address per crypto peer for proper operation. PAT works by masquerading multiple crypto peers behind a single IP address.

The IPsec NAT Traversal feature (NAT-T) introduces support for IPsec traffic to travel through NAT or PAT devices by encapsulating both the IPsec SA and the ISAKMP traffic in a UDP wrapper. NAT-T was first introduced in Cisco IOS version 12.2(13)T and is auto-detected by VPN devices. There are no configuration steps for a Cisco IOS router running this release or later because it is enabled by default as a global command. The NAT-T feature detects a PAT device between the crypto peers and negotiates NAT-T if it is present.

For more information on IPsec NAT-T (also known as transparency), see the following URL: http://www.cisco.com/en/US/docs/ios/12_2t/12_2t13/feature/guide/ftipsnat.html.

DMVPN designs are compatible with NAT-T in a hub-and-spoke deployment model. Cisco IOS Release 12.3(9a) or 12.3(12)T2 or later is required on both the headend and branch routers to support DMVPN topologies where the headend router is behind NAT.

In spoke-to-spoke designs, NAT is a greater concern. NHRP registrations/resolutions do not allow spoke-to-spoke crypto sessions to properly form if a spoke is behind a NAT device. In this situation, the IPsec SA does not establish, and the branch continues to send packets to the remote branch via the hub. Depending on the software version, these packets may be process-switched. Because of this, spoke-to-spoke topologies should be avoided when spokes are behind NAT boundaries. This restriction was removed in IOS 12.4(6)T and later versions.

In dual DMVPN cloud topologies, regardless of deployment model, the Single Tier Headend Architecture can be configured with IPsec in tunnel mode. IPsec transport mode is recommended for DMVPN designs and required when NAT/PAT is involved, because of the following:

- When the branch sends its NHRP registrations to the headend, the headend sees both the branch outside NAT address and inside host GRE address. The headend selects the outside NAT address from the branch for its use.

- Different branch routers can use the same (overlapping) inside host GRE address, because the outside NAT address is unique.

The Dual Tier Headend Architecture (not recommended now) requires the IPsec tunnels to be configured in tunnel mode. Note the following caveats with this configuration:

- The headend router can see only the inside host GRE address of the branch router in the NHRP registrations coming from the branch router.
- Branch routers must therefore have unique inside host GRE address, requiring coordination for all branches in the DMVPN cloud.

Note As a caveat, IPsec tunnels generated via DMVPN between headend and branch are not supported with the headend behind NAT/PAT when using the Cisco Catalyst 6500 or Cisco 7600 with VPNSM. If this is a design requirement, Cisco recommends using the Cisco 7200VXR or other Cisco IOS router at the headend.

In the event that there is more than one branch router behind a single NAT device, DMVPN can only support this configuration if the NAT device translates each branch router to a unique outside NAT IP address.

Firewall Considerations

This section discusses the various firewall considerations when implementing a DMVPN design.

Headend or Branch

Depending on the crypto and DMVPN headend or branch placements, the following protocols and ports are required to be allowed:

- UDP Port 500—ISAKMP as source and destination
- UDP Port 4500—NAT-T as a destination
- IP Protocol 50—ESP
- IP Protocol 51—AH (if AH is implemented) (this is not recommended)

Network location of the crypto headend in relation to the headend firewall(s) impacts both the accessibility and performance of both systems. The network manager must ensure that all firewalls are properly configured to allow the tunnel traffic bi-directionally. The crypto headend must be accessible to the branch router because all crypto sessions are initiated by the branch router.

Crypto Access Check

DMVPN may use tunnel profiles, which eliminate the need to statically define crypto ACLs that would normally be used to match packets that require encryption. However, this functionality is still handled by the software dynamically. The **show crypto map** Cisco IOS command can be used to view the access list that is generated. The access list matches the GRE packet from the tunnel endpoints. If the encryption and the GRE tunnel are not using the same endpoints, tunnel profiles cannot be used, and crypto access lists or dynamic crypto maps must be configured. The access list should reflect the crypto endpoints and not the GRE endpoints. This configuration is required when implementing a Dual Tier Headend Architecture.

Common Configuration Mistakes

The following sections discuss common mistakes and problems encountered when configuring DMVPN.

Advertising Tunnel Endpoints in the Routing Protocol

It is possible to include the tunnel endpoints in the internal routing protocol. In this case, the headend address is typically a globally-routed address on the public Internet. Although not a best practice, this address may also be advertised internally. In this situation, the DMVPN subnet initially works properly, but after routing has converged, the branch may attempt to reach the headend public address via the DMVPN tunnel rather than the outside interface. This causes the tunnel to fail and consequently the routing protocol to stop advertising the public address, allowing the DMVPN tunnel to once again function correctly. The process repeats continually, and is sometimes referred to as “recursive routing or forwarding”.

IPsec Transform Set Matches

At least one matching IPsec transform set must be configured between two crypto peers. When specifying a particular strength of encryption algorithm, a common strength encryption algorithm and transform set should also be configured on both the headend and branches. Failure to do so prevents the IPsec tunnel from starting.

ISAKMP Policy Matching

There is a default ISAKMP policy present in all Cisco IOS devices. This default is encryption DES, HMAC of SHA, IKE Authentication of RSA signature, and DH group 1. If a stronger ISAKMP policy is desired, both sides must support the policy.

It is common, but not required, to use the same encryption level transform set and hash methods in the ISAKMP policy and IPsec transform set.