

Application Note: AoE WAN Connectivity

Abstract

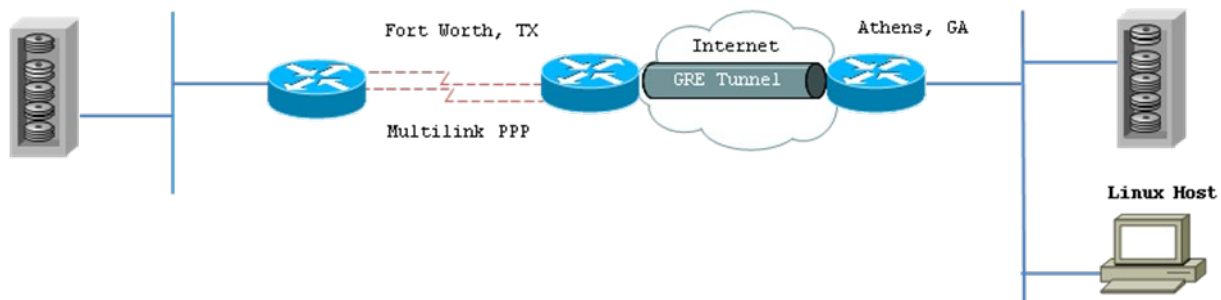
During a recent customer Proof of Concept (POC), Coraid Engineering has identified methods for AoE WAN connectivity and remote management of AoE disk shelves.

POC Network Topology

The Athens side of the POC consists of a Cisco 2811 router with a Gigabit Ethernet interface connected to an EtherDrive® storage appliance and a serial interface connected to a DS1 (T1) Internet connection.

The Fort Worth peer is a Cisco 2811 router with a DSL connection to the Internet and two serial interfaces linking to a Cisco 1751 router with a FastEthernet interface connecting an EtherDrive® storage appliance.

Figure 1 – Proof of Concept Network Topology



GRE Tunneling and Transparent Bridging

Because the customer requires AoE access over the public Internet, a GRE tunnel was configured to link the sites. To provide AoE connectivity, transparent bridging was configured on each interface in the path from initiator to target.

GRE Tunnel

Generic Route Encapsulation (GRE) (ethertype 0x800) is a technology frequently used to transport non-TCP/IP protocols across an IP network. GRE encapsulates an arbitrary payload in an IP packet. The foreign protocol payload in this case is AoE (ethertype 0x88A2). Example 1 shows the GRE configuration of two Internet-connected Cisco 2811 routers.

Note: Private (RFC 1918) IP addresses are used in all examples.

Example 1 – Base GRE Configuration

Athens Site

```
hostname Coraid-2811
!
interface FastEthernet0/0
  description Internet
  ip address 10.10.1.1 255.255.255.0
!
interface FastEthernet0/1
  description Storage VLAN
  ip address 172.16.1.1 255.255.255.0
!
interface Tunnel0           ! Tunnel0 is a logical interface
description AoE Tunnel
no ip address
tunnel source 10.10.1.1     ! Internet-facing interface FE0/0
tunnel destination 10.20.1.1 ! IP address of GRE Peer
```

Fort Worth Site

```
hostname Customer-2811
!
interface FastEthernet0/0
  description Internet
  ip address 10.20.1.1 255.255.255.0
!
interface FastEthernet0/1
  description Storage VLAN
  ip address 172.16.2.1 255.255.255.0
!
interface Tunnel0
description AoE Tunnel
no ip address
tunnel source 10.20.1.1
tunnel destination 10.10.1.1
```

GRE Security Concerns

It is important to understand that data is only *encapsulated* in a GRE tunnel. Because AoE data is sent over the tunnel in clear text, it is vulnerable to interception at any hop from source to destination over the Internet. IPSec should be used to ensure confidentiality and integrity of the AoE payload.

Configure IPSec Over GRE

Example 2 updates the GRE tunnel with IPSec encryption and integrity validation. The example includes 128-bit AES encryption which is sufficient for most commercial applications. AES 256-bit should be used for data subject to regulatory compliance.

Example 2 – IPSec Over GRE

Athens Site

```
hostname Coraid-2811
!
ip access-list extended AoE
  permit ip host 10.10.1.1 host 10.20.1.1
  ! Select traffic for IPSec. Only traffic between the two tunnel
  ! endpoints is encrypted.
  !
crypto isakmp policy 1           ! Define IKE Phase I Policy Set
  encr aes                       ! Use 128-bit AES Encryption
  authentication pre-share      ! Authenticate peer w/pre-shared key
  group 5                       ! Use 1536-bit prime for DH exchange
  ! In IKE Phase I, peers exchange policy, perform a Diffie-Hellman
  ! key exchange, and authenticate their peer with a pre-shared key.
  ! The resulting secure tunnel is used to protect the policy exchange
  ! for IKE Phase II.
crypto isakmp key EtherDrive address 10.20.1.1 no-xauth
  ! Define the pre-shared secret string "EtherDrive" to authenticate
  ! peer without additional username/password exchange "no-xauth".
  !
crypto ipsec transform-set CORAID esp-aes esp-sha-hmac
  mode transport
  ! The transport set defines the algorithms used for encryption
  ! and packet integrity. Because this is a site to site connection
  ! between two IP addresses (peers), transport mode is used and avoids
  ! the additional encapsulation penalty of a second IP header .
  !
```

Example 2 – (continued)

```
crypto map AoE-WAN 1 ipsec-isakmp ! IKE Phase II Policy Set
  set peer 10.20.1.1 ! Define crypto peer
  set transform-set CORAID ! Use transform-set CORAID
  set pfs group5 ! Use perfect forward secrecy
  match address AoE ! Use access-list AoE
! The crypto map defines how the data is encrypted and authenticated.
! Perfect Forward Secrecy demands that a full Diffie-Hellman key
! exchange is performed at the end of the IKE Phase II Security
! Association lifetime (3600 seconds on a Cisco router).
!
interface FastEthernet0/0
  description Internet
  ip address 10.10.1.1 255.255.255.0
  crypto map AoE-WAN ! Bind crypto map to interface
!
interface FastEthernet0/1
  description Storage VLAN
  ip address 172.16.1.1 255.255.255.0
!
interface Tunnel0
  description AoE Tunnel
  no ip address
  tunnel source 10.10.1.1
  tunnel destination 10.20.1.1
```

Fort Worth Site

```
hostname Customer-2811
!
ip access-list extended AoE
  permit ip host 10.20.1.1 host 10.10.1.1
!
crypto isakmp policy 1
  encr aes
  authentication pre-share
  group 5
crypto isakmp key EtherDrive address 10.10.1.1 no-xauth
!
crypto ipsec transform-set CORAID esp-aes esp-sha-hmac
  mode transport
!
```

Example 2 – (continued)

```
crypto map AoE-WAN 1 ipsec-isakmp
  set peer 10.10.1.1
  set transform-set CORAID
  set pfs group5
  match address AoE
!
interface FastEthernet0/0
  description Internet
  ip address 10.20.1.1 255.255.255.0
  crypto map AoE-WAN
!
interface FastEthernet0/1
  description Storage VLAN
  ip address 172.16.2.1 255.255.255.0
!
interface Tunnel0
  description AoE Tunnel
  no ip address
  tunnel source 10.20.1.1
  tunnel destination 10.10.1.1
```

Transparent Bridging

AoE operates at layer 2 of the OSI model and it is not routable. The solution for AoE WAN transport is to employ transparent bridging (IEEE 802.1d). On links that support TCP/IP *and* bridging, IP packets are routed based on destination IP address and AoE packets are forwarded based on destination MAC addresses. Although TCP/IP and AoE can share a WAN connection, this solution should only be used for remote management of AoE appliances. Dedicated bandwidth should be provisioned for data replication.

Example 3 illustrates the simple addition to the GRE configuration that enables AoE bridging across the tunnel interface.

Note: It is un-necessary to configure a bridge-group on the interface connecting to the ISP. Interface FastEthernet0/0 is merely the GRE tunnel endpoint.

Example 3 – Add Transparent Bridging to IPsec Over GRE Tunnel

Athens Site

```
hostname Coraid-2811
!
ip access-list extended AoE
  permit ip host 10.10.1.1 host 10.20.1.1
!
crypto isakmp policy 1
  encr aes
  authentication pre-share
  group 5
crypto isakmp key EtherDrive address 10.20.1.1 no-xauth
crypto ipsec transform-set CORAID esp-aes esp-sha-hmac
  mode transport
!
crypto map AoE-WAN 1 ipsec-isakmp
  set peer 10.20.1.1
  set transform-set CORAID
  set pfs group5
  match address AoE
!
interface FastEthernet0/0
  description Internet
  ip address 10.10.1.1 255.255.255.0
!
interface FastEthernet0/1
  description Storage VLAN
  ip address 172.16.1.1 255.255.255.0
  bridge-group 1           ! Add interface to bridge-group 1
!
interface Tunnel0
  description AoE Tunnel
  no ip address
  tunnel source 10.10.1.1
  tunnel destination 10.20.1.1
  crypto map AoE-WAN
  bridge-group 1           ! Add interface to bridge-group 1
!
bridge 1 protocol ieee
! Define bridge-group 1 using 802.1d spanning-tree protocol. All
! interfaces that need to bridge AoE must be a member of this
! bridge-group. There must be at least two members of any given
! bridge-group. FastEthernet0/0 does not need to be a member of the
! bridge-group since its role is simply a tunnel endpoint.
```

Example 3 – (continued)

Fort Worth Site

```
hostname Customer-2811
!
ip access-list extended AoE
 permit ip host 10.20.1.1 host 10.10.1.1
!
crypto isakmp policy 1
 encr aes
 authentication pre-share
 group 5
crypto isakmp key EtherDrive address 10.10.1.1 no-xauth
!
crypto ipsec transform-set CORAID esp-aes esp-sha-hmac
 mode transport
!
crypto map AoE-WAN 1 ipsec-isakmp
 set peer 10.10.1.1
 set transform-set CORAID
 set pfs group5
 match address AoE
!
interface FastEthernet0/0
 description Internet
 ip address 10.20.1.1 255.255.255.0
!
interface FastEthernet0/1
 description Storage VLAN
 ip address 172.16.2.1 255.255.255.0
 bridge-group 1
!
interface Tunnel0
 description AoE Tunnel
 no ip address
 tunnel source 10.20.1.1
 tunnel destination 10.10.1.1
 crypto map AoE-WAN
 bridge-group 1
```

Transparent Bridging Over Dedicated Circuits

The POC network also allowed Coraid to demonstrate AoE connectivity over dedicated (leased-line) circuits. In figure 1, the Fort Worth end of the connection separates the router terminating the GRE tunnel

and the EtherDrive® appliance by serial links. DTE/DCE cables were used to simulate leased line connections.

PPP WAN Encapsulation

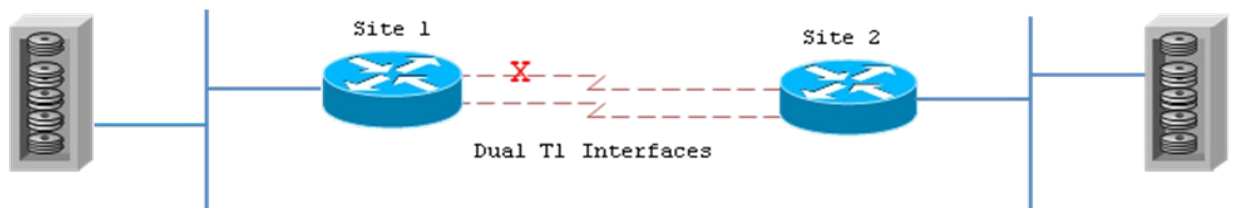
PPP (Point-to-Point protocol) is a popular WAN encapsulation that allows transport encapsulation for TCP/IP, IPX and AppleTalk. PPP also supports remote bridging for layer 2 protocols.

Spanning-Tree Considerations

The Spanning Tree protocol is a very important component of 802.1d bridging. Spanning Tree allows for multiple layer 2 links between sites for redundancy and ensures a loop-free topology. Without Spanning Tree, multiple links cannot be employed within a bridge group without causing a bridge loop and loss of connectivity.

The drawback of Spanning Tree is that only one bridge link in a cluster can be in a forwarding state. Figure 2 shows a scenario where two PPP connections connect remote sites for greater bandwidth, but Spanning Tree allows only one circuit to forward layer 2 traffic. If the serial lines are configured to route TCP/IP and bridge AoE, a routing protocol could load balance across both links, whereas AoE traffic would be restricted to the interface in forwarding state.

Figure 2 – Spanning Tree Loop Prevention



Multilink PPP

Figure 3 illustrates how a Multilink interface combines multiple physical PPP links into a single virtual interface. Because Spanning Tree perceives a single interface, all links forward traffic.

Figure 3 – Multilink PPP Interface Aggregation



Example 4 demonstrates the additional configuration to the Fort Worth routers to create a multilink PPP interface from two physical serial interfaces and enable bridging. For this example, the storage appliance was moved behind the Cisco 1751 router shown in Figure 1. The IPSec configuration is omitted for clarity.

Example 4 – Configure Bridging Over Multilink PPP

Fort Worth 2811

```
hostname Customer-2811
!
interface FastEthernet0/0
  description Internet
  ip address 20.20.1.1 255.255.255.0
!
interface Tunnel0
  description AoE Tunnel
  no ip address
  tunnel source 10.20.1.1
  tunnel destination 10.10.1.1
  bridge-group 1
!
interface Multilink1      ! Multilink1 is a logical interface
  mtu 9344                ! Set interface MTU to allow 9K frames
  ip address 192.168.1.1 255.255.255.252
  ppp multilink          ! Enable multilink on the interface
  ppp multilink group 1  ! Define multilink group number
  bridge-group 1        ! Add multilink interface to bridge-group
!
interface Serial0/0/0
  mtu 9344                ! Set interface MTU to allow 9K frames
  no ip address          ! Aggregated interface does not need IP
  encapsulation ppp      ! Interface must use PPP to multilink
  clock rate 4000000     ! With DTE/DCE cable, DCE provides clocking
  ppp multilink          ! Enable multilink on the interface
  ppp multilink group 1 ! Define multilink group number
```

Example 4 – (continued)

```
interface Serial0/0/1
  mtu 9344                ! Set interface MTU to allow 9K frames
  no ip address          ! Aggregated interface does not need IP
  encapsulation ppp      ! Interface must use PPP to multilink
  clock rate 4000000     ! With DTE/DCE cable, DCE provides clocking
  ppp multilink          ! Enable multilink on the interface
  ppp multilink group 1 ! Define multilink group number
```

```

!
bridge 1 protocol ieee

Fort Worth 1751
hostname Customer-1751
!
interface Multilink1
  mtu 9344
  ip address 192.168.1.2 255.255.255.252
  ppp multilink
  ppp multilink group 1
  bridge-group 1
!
interface FastEthernet0/0
ip address 192.168.2.1 255.255.255.0
bridge-group 1
!
interface Serial1/0
  mtu 9344
  no ip address
  encapsulation ppp
  ppp multilink
  ppp multilink group 1
!
interface Serial1/1
  mtu 9344
  no ip address
  encapsulation ppp
  ppp multilink
  ppp multilink group 1
!
bridge 1 protocol ieee

```

HDLC, Frame Relay & X.25

Although HDLC, Frame Relay, and X.25 encapsulations allow for configurable interface MTU and bridged AoE connectivity, they do not have multilink interface capability. In the case of Frame Relay and X.25, the carrier network may not support 9K frames even if it is configured on the serial interface.

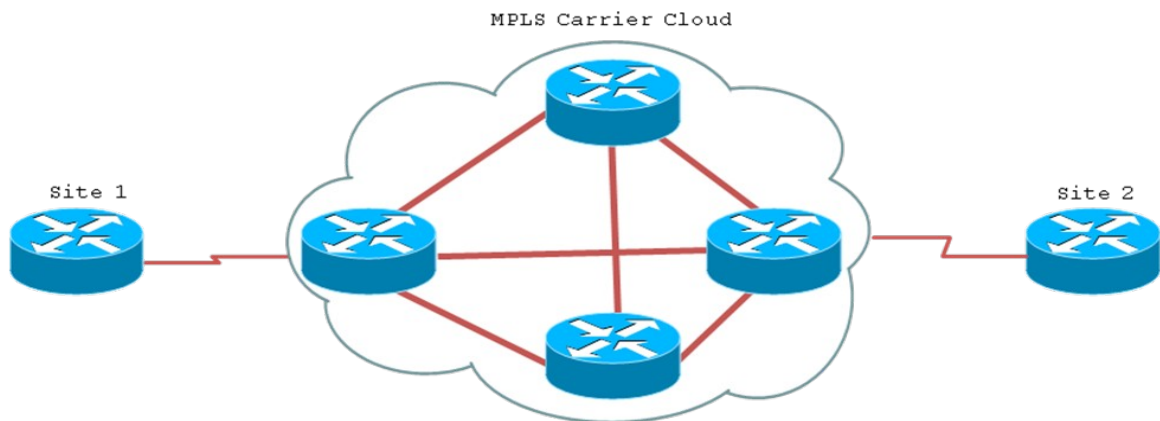
If multiple serial interfaces using HDLC, Frame Relay, or X.25 encapsulation are provisioned to a site, Spanning Tree will forward on only one interface and place additional interfaces into blocking state to ensure a loop-free layer 2 topology.

Spanning Tree blocking only affects bridged traffic. Layer 3 traffic can be routed over multiple links even if Spanning Tree has blocked a link at layer 2.

MPLS

Major Service Providers such as AT&T have warmly embraced Multiprotocol Label Switching (MPLS) in their networks. Two important benefits of adopting MPLS within the carrier network are Traffic Engineering (improves reliability), and the freedom to replace BGP routing with MPLS in their backbone networks. The elimination of BGP significantly reduces the memory and CPU requirements of backbone routers. Figure 4 illustrates a simple MPLS topology.

Figure 4 – Site to Site WAN Over MPLS



In many cases, customers benefit from a sizable reduction in monthly recurring fees for WAN circuits, and increased reliability. Example 5 demonstrates the edge router configurations of two sites connected by Frame Mode MPLS.

Unfortunately, transparent bridging cannot be applied to an interface configured for MPLS (regardless of encapsulation). The solution is to employ VPLS (Virtual Private LAN Services) that run over the SP (Service Provider) MPLS network. At the CE (customer edge) router, a simple 802.1q Ethernet trunk is presented to the SP edge device. VPLS is currently an IETF Draft Standard. Check with your Service Provider for availability.

Troubleshooting

All examples assume Cisco routers and switches. For other vendors, please refer to the appropriate documentation for guidance to perform these procedures.

Troubleshooting GRE

GRE configurations are relatively straightforward when used solely to transport bridged traffic. All you need is a definition for the tunnel source and the tunnel destination. Since the configuration in example 1 does not use an IP address on the Tunnel0, IP routing is not possible over the GRE Tunnel. This is desirable if the tunnel interface is dedicated to AoE traffic.

The most common reasons for the tunnel interface to remain down are IP routing and misconfigured tunnel source or destination. When creating a GRE tunnel over the Internet, the source and destination addresses of the tunnel must be globally routable IP addresses (RFC1918 private addresses will not work) and they must be reachable from each peer.

```
Customer-2811# show interface tunnel0
Tunnel0 is up, line protocol is up
  Hardware is Tunnel
  MTU 9108 bytes, BW 100 Kbit/sec, DLY 50000 usec,
    reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel source 172.16.222.11, destination 172.16.113.9
  Tunnel protocol/transport GRE/IP
```

Troubleshooting Bridging

In the context of this Application Note, the configuration of transparent bridging consists of just two commands. One is a global command and the other an interface command:

- **bridge 1 protocol ieee**
- **bridge-group 1**

The first command is global and creates bridge-group 1 and enables IEEE 802.1d Spanning Tree protocol. Other spanning tree protocols include DEC and IBM. The DEC protocol is largely obsolete and IBM is only required for Source Route Bridging in IBM mainframe environments.

The second command causes an interface to join bridge-group 1 and forward packets based on their layer 2 MAC address. All interfaces in the path from the AoE Initiator to the AoE target must be members of the bridge-group.

In some cases, you may wish to restrict the hosts that are allowed to transmit over the bridged links. The following global command illustrates how to manually restrict bridged traffic.

Please note that you need to configure at least two MAC addresses allowed to forward over the bridge. This MAC filter should be configured on both ends of serial or GRE links to prevent unauthorized hosts from using that link.

- **bridge 1 address 00c1.10fb.3365 forward**

Spanning-tree

The Spanning Tree protocol (STP) is an essential element in bridged networks. Each bridge helps to build a loop-free topology by sending periodic BPDU (Bridge Protocol Data Units) to the well-known multicast address 01:80:C2:00:00:00. Through BPDU's, the bridge finds duplicate paths to a destination and blocks them.

When a bridge initializes, there is always the risk of a loop present. Because of this, no member of the bridge-group is allowed to transmit until STP builds its topology table based on receipt of BPDUs from peer bridges.

All router interfaces go through a three-stage process before forwarding:

- **Listening** – The interface listens for BPDUs from other bridges to build its topology.
- **Learning** – The interface begins building a forwarding table based on MAC addresses.
- **Forwarding** – After about 30 seconds, an interface will transition to a forwarding state if it is not put into blocking state by STP.

The **show spanning-tree** command is useful to determine the bridging state interfaces are in.

```
Customer-2811# show spanning-tree
```

```
Bridge group 1 is executing the ieee compatible Spanning Tree protocol  
Bridge Identifier has priority 1, address 0014.a925.7b29  
Configured hello time 2, max age 20, forward delay 15  
We are the root of the spanning tree  
Topology change flag not set, detected flag not set  
Number of topology changes 5 last change occurred 00:47:27 ago  
Times: hold 1, topology change 35, notification 2  
hello 2, max age 20, forward delay 15  
Timers: hello 1, topology change 0, notification 0, aging 300  
  
Port 5 (FastEthernet0/1) of Bridge group 1 is forwarding
```

```
Port path cost 19, Port priority 128, Port Identifier 128.5.  
Designated root has priority 1, address 0014.a925.7b29  
Designated bridge has priority 1, address 0014.a925.7b29  
Designated port id is 128.5, designated path cost 0  
Timers: message age 0, forward delay 0, hold 0  
Number of transitions to forwarding state: 1  
BPDU: sent 6792, received 4
```

Port 15 (Tunnel0) of Bridge group 1 is **forwarding**

```
Port path cost 45575, Port priority 128, Port Identifier 128.15.  
Designated root has priority 1, address 0014.a925.7b29  
Designated bridge has priority 1, address 0014.a925.7b29  
Designated port id is 128.15, designated path cost 0  
Timers: message age 0, forward delay 0, hold 0  
Number of transitions to forwarding state: 1  
BPDU: sent 6787, received 0
```

Port 16 (Multilink1) of Bridge group 1 is **forwarding**

```
Port path cost 323, Port priority 128, Port Identifier 128.16.  
Designated root has priority 1, address 0014.a925.7b29  
Designated bridge has priority 1, address 0014.a925.7b29  
Designated port id is 128.16, designated path cost 0  
Timers: message age 0, forward delay 0, hold 0  
Number of transitions to forwarding state: 1  
BPDU: sent 1432, received 5362
```

Root Bridge

In for every STP bridge-group, one bridge is elected as the “root bridge”. The bridge with the lowest Bridge ID is selected as root. It is good practice to manually place the root bridge at the “top” of the bridged network. It would be most appropriate to make a router in the data center the root bridge. Since the Bridge ID is a combination of MAC address and Priority, the administrator can force a router to become the root bridge.

```
Customer-2811(config)# bridge 1 priority 1  
Customer-2811# show spanning-tree
```

```
Bridge group 1 is executing the ieee compatible Spanning Tree  
protocol
```

```
Bridge Identifier has priority 1, address 0014.a925.7b29  
Configured hello time 2, max age 20, forward delay 15  
We are the root of the spanning tree
```

Troubleshooting IPsec

IPsec can be very intimidating to configure and troubleshoot. Fortunately, the IPsec elements required to encrypt the GRE tunnel are about as straightforward as you will see in an IPsec design.

One of the most important things to determine is whether the peer is reachable via the GRE tunnel *before* IPsec is configured. You can waste a lot of time chasing a perceived IPsec issue when in fact, the GRE tunnel is misconfigured.

Useful Show Commands

The **show crypto isakmp policy** command allows you to verify the IKE Phase I policy. Each IPsec peer must have at least one matching policy for the Phase I exchange to complete. Verify both peers have a complementary policy. The only parameter that does not have to match is the lifetime.

The “Default protection suite” is a system default. It is never used unless no other policy is configured.

```
Customer-2811# show crypto isakmp policy
```

```
Global IKE policy
```

```
Protection suite of priority 1
```

```
  encryption algorithm:  AES - Advanced Encryption Standard (128 bit keys).
```

```
  hash algorithm:       Secure Hash Standard
```

```
  authentication method: Pre-Shared Key
```

```
  Diffie-Hellman group: #5 (1536 bit)
```

```
  lifetime:             86400 seconds, no volume limit
```

```
Default protection suite
```

```
  encryption algorithm:  DES - Data Encryption Standard (56 bit keys).
```

```
  hash algorithm:       Secure Hash Standard
```

```
  authentication method: Rivest-Shamir-Adleman Signature
```

```
  Diffie-Hellman group: #1 (768 bit)
```

```
  lifetime:             86400 seconds, no volume limit
```

Use the **show crypto ipsec transform** command to verify that both peers are using identical encryption and authentication algorithms. The mode (transport in our case) must also match.

```
Customer-2811# show crypto ipsec transform-set
```

```
Transform set CORAID: { esp-aes esp-sha-hmac }
```

```
will negotiate = { Transport, },
```

The **show crypto map** command allows you to validate the IKE Phase II configuration. In the output, be sure all fields have a value. If the output shows peer, access-list, transform-set or interface missing, the connection will fail. The names of the access-lists and transform-set are case-sensitive. That is one of the most difficult configuration mistakes to find.

```
Customer-2811# show crypto map
Crypto Map "AoE-WAN" 1 ipsec-isakmp
  Peer = 12.51.113.9
  Extended IP access list AoE
    access-list AoE permit ip host 66.182.222.11 host 12.51.113.9
  Current peer: 12.51.113.9
  Security association lifetime: 4608000 kilobytes/3600 seconds
  PFS (Y/N): Y
  DH group: group5
  Transform sets={
    CORAID,
  }
  Interfaces using crypto map AoE-WAN:
    FastEthernet0/0
```

References

The following resources will be useful to customers who decide to configure AoE Over WAN. All links are hot (clickable).

Online

Coraid Resource Center

<http://www.coraid.com/s.nl/sc.12/.f>

Cisco Legacy Product Documentation

<http://www.cisco.com/univercd/home/home.htm>

Cisco Current Product Documentation

<http://www.cisco.com/web/psa/products/index.html>

Cisco GRE Configuration

http://www.cisco.com/en/US/docs/ios/12_2/interface/configuration/guide/icflogin.html#wp1012601

Cisco IOS Bridging Command Reference

http://www.cisco.com/en/US/docs/ios/bridging/command/reference/br_a1.html#wp10

Cisco IOS Command References

http://www.cisco.com/en/US/products/ps6350/prod_command_reference_list.html

IPSec Over GRE Configuration Guide

http://www.cisco.com/univercd/cc/td/doc/product/vpn/solution/aswan15/sig/sig_05.htm

Spanning Tree Protocol

http://en.wikipedia.org/wiki/BPDU#Bridge_Protocol_Data_Units_.28BPDU.29

VPLS Technology Presentations

http://www.cisco.com/en/US/products/ps6648/products_ios_protocol_option_home.html

The MPLS-VPLS Resource Center

<http://www.mplsrc.com>

Books

IPSEC, 2nd Edition

<http://www.bookpool.com/sm/013046189X>

Building Cisco Multilayer Switched Networks (BCMSN)

<http://www.bookpool.com/sm/1587052733>

Network Design and Case Studies (CCIE Fundamentals), 2nd Edition

<http://www.ciscopress.com/bookstore/product.asp?isbn=1578701678>