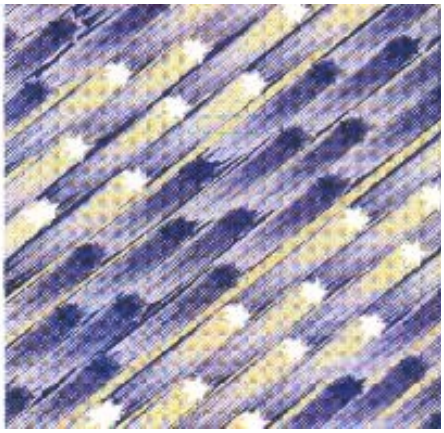


Fault Tolerant IP Security In a Frame Relay Network

Matt Trzyna

IBM Networking Hardware Division
Product Installation Support
Cary, North Carolina

Abstract

This document describes the Proof of Concept testing done to develop and validate key functional pieces of a high level network design for a large router based network. The main functional components of the design that will be discussed in this paper are:

- Frame Relay
- IP
- OSPF with multiple areas and address summarization
- IP Security (IPSec)
- Virtual Router Redundancy Protocol (VRRP)

Table of Contents

Abstract . . . . . Page 2
Preface . . . . . Page 4
Network Design Overview . . . . . Page 4
Testing Overview . . . . . Page 4
Document Overview . . . . . Page 5
Network Design Details . . . . . Page 7
Network Objectives/Requirements: . . . . . Page 7
Functional Design Details . . . . . Page 7
Initial Frame Relay Design . . . . . Page 8
Initial IP/OSPF and VRRP Design . . . . . Page 9
Initial IPSec Design . . . . . Page 11
Setup, Implementation, and Testing . . . . . Page 13
Test Overview . . . . . Page 13
Frame Relay and IP/OSPF Infrastructure Implementation/Testing . . . . . Page 14
Initial Frame Relay and IP/OSPF Configuration . . . . . Page 14
Frame Relay Configuration . . . . . Page 15
IP Configuration . . . . . Page 16
OSPF Configuration . . . . . Page 16
Test Procedure Overview . . . . . Page 19
Initial Configuration Test Results . . . . . Page 21
OSPF Configuration Redesign . . . . . Page 23
OSPF Redesign Overview . . . . . Page 23
OSPF Redesign Test Results . . . . . Page 25
VRRP Implementation/Testing . . . . . Page 28
VRRP Configuration . . . . . Page 28
VRRP Configuration Testing Results . . . . . Page 29
IPSec Design, Implementation, and Testing . . . . . Page 35
IPSec Output/Input Processing . . . . . Page 35
IPSec Design . . . . . Page 37
IPSec Availability Challenge . . . . . Page 38
IPSec Configuration Details . . . . . Page 39
Tunnel Configuration Details . . . . . Page 39
Access Control Configuration Details . . . . . Page 41
IPSec Configuration Testing Results . . . . . Page 48
Summary . . . . . Page 50
Appendix A: Test Procedures . . . . . Page 51
Appendix B: Recommended Reading . . . . . Page 55

Preface

This document describes the findings of Proof of Concept testing done to validate key functional components of a proposed high level network design.

Network Design Proposal Overview

Design Objective

The main objective of the design was to replace two separate leased line SNA and X.25 based IP networks with a common, multi-protocol, partially meshed public Frame Relay infrastructure based on IBM 2216 and 2210 routers.

Comonents/Requirements

Major components/requirements of the design were:

- DLSw for SNA traffic
- Multiple OSPF areas with address summarization for limiting routing table size and routing protocol traffic
- Security: privacy and integrity of the user and application data flowing between sites
- High availability
  - no single point of failure for overall network operation
  - redundant Frame Relay PVC paths for all routers
  - WAN backup dial out capability
  - maintenance of IPSec functionality in all failure scenarios
  - Ease of configuration
- Router CPU utilization efficiency

Testing Overview

The overall objective of the testing was to develop implementation recommendations and demonstrate that the proposed network design would meet the customer's key requirements. This was accomplished by implementing a small test network that could exercise the following design elements:

- Frame Relay
- OSPF
- IPSec
- High availability
- Ease of configuration
- Router CPU efficiency

DLSw and dial backup were not included in the testing due to scheduling constraints.

Following is a list of the equipment and associated code levels used in the test:

- Networking Equipment
  - Routers: The GUI configuration tool was used to configure all the routers.
    - IBM 2210 24T: MRS 3.2
    - IBM 2216 M400: MAS 3.2
  - LAN Switches (for token ring infrastructure): IBM 8270 M800
  - Frame Relay Switch: IBM 2225 M400
- Work Stations: IBM IntelliStations, NT

A copy of the test cases which were run to validate the test system can be found in *Appendix A* (page 51).

Document Overview

The objective of this document is to share implementation and testing experiences in the hope that the reader can save time and effort in the future by:

- using some of the recommended techniques when similar situations are encountered.
- avoiding some of the problems encountered during the test.

Though a shorter paper would have resulted if only the final configuration was described, I felt greater benefit would be provided by describing the problems uncovered and the design and implementation solutions developed to resolve them.

The paper is organized as follows:

- Network Design Details
- Proof of Concept Setup, Implementation, and Test
  - Frame Relay
  - IP/OSPF
  - VRRP
  - IPSec
- Summary

A basic understanding of the product functions discussed is assumed as explanations of their operation and specific equipment configuration details are only provided to explain key concepts. All recommendations are based on the implementation characteristics of the equipment listed above. These characteristics may not be the same as products from other vendors. Finally, the recommendations made in this document should not be viewed as the only possible methods for achieving the design's objectives. The complexity of some of the functions naturally implies that there could be other valid solutions.

Finally, it is strongly advised to view the diagrams in color as use of color coding was the only way to display some of the key details.

Send any questions or comments to:

Matt Trzyna
trzyna@us.ibm.com

Network Design Details

Network Objectives/Requirements:

A network design was needed to replace two separate leased line SNA and X.25 based IP networks with a single network. Key requirements for the new network were:

- **Multiprotocol Support**: SNA and IP traffic must be carried by a common WAN infrastructure.
- **Connectivity**: All branch sites must be able to communicate with each other as well as with centrally located servers and mainframes.
- **Availability:** There must be no single point of network infrastructure failure which would inhibit sites from accessing each other or the centrally located mainframes and servers.
- **Security**: Privacy and integrity must be provided for data flowing between clients and servers located in branch offices as well as between clients in branch offices and the centrally located mainframes and servers.
- **Performance**: Though no specific application level performance requirements were identified, due to the size of the network steps had to be taken to minimize router CPU utilization and limit overhead WAN traffic as much as possible.
- **Ease of Configuration**: Due to the large number of sites involved, the addition of branch offices to the network must be as simple and standardized as possible.

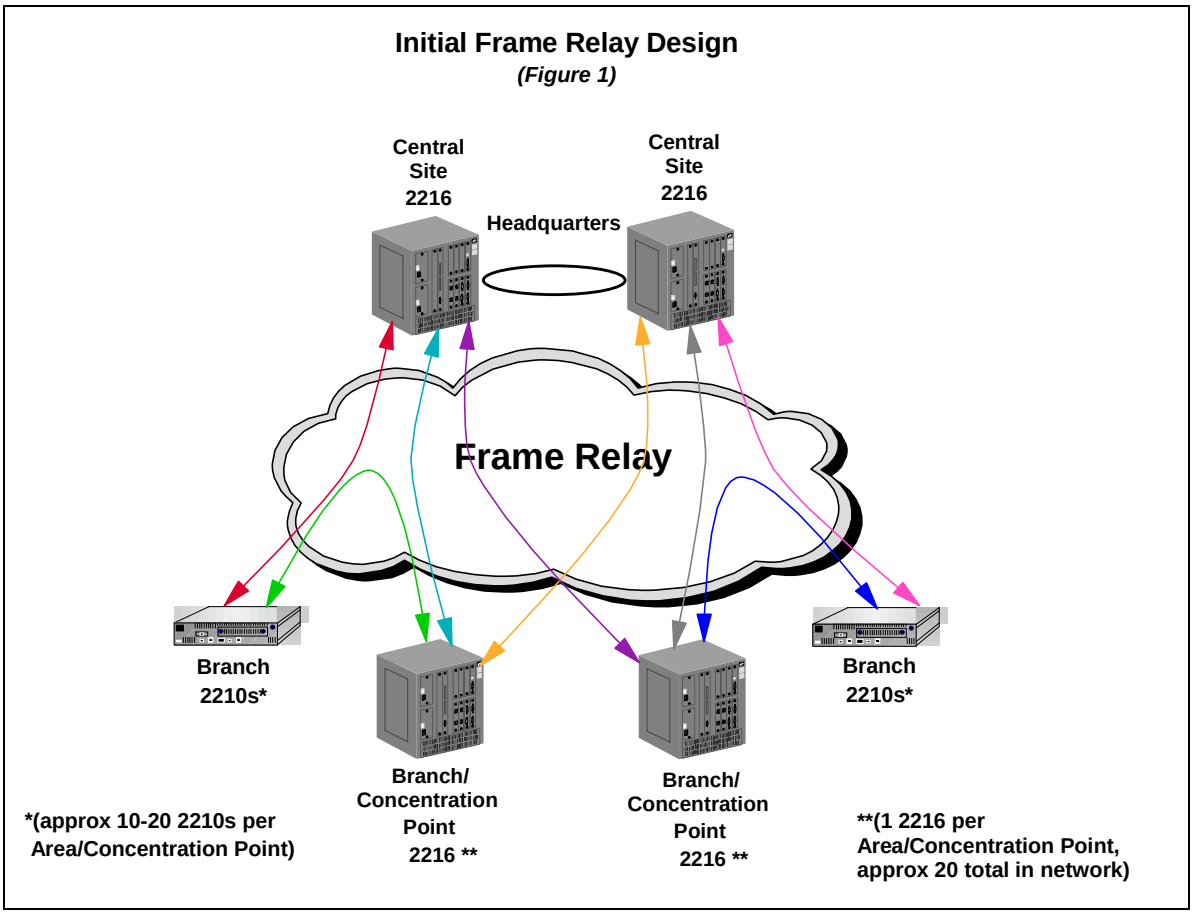
A solution based on IBM 2210 and 2216 routers was chosen to meet these requirements.

Functional Design Details

Due to the difficulty in clearly showing all facets of a router based network in one diagram, the description will be divided into three distinct design components:

- Frame Relay
- IP, OSPF, and VRRP
- IPSec

Initial Frame Relay Design



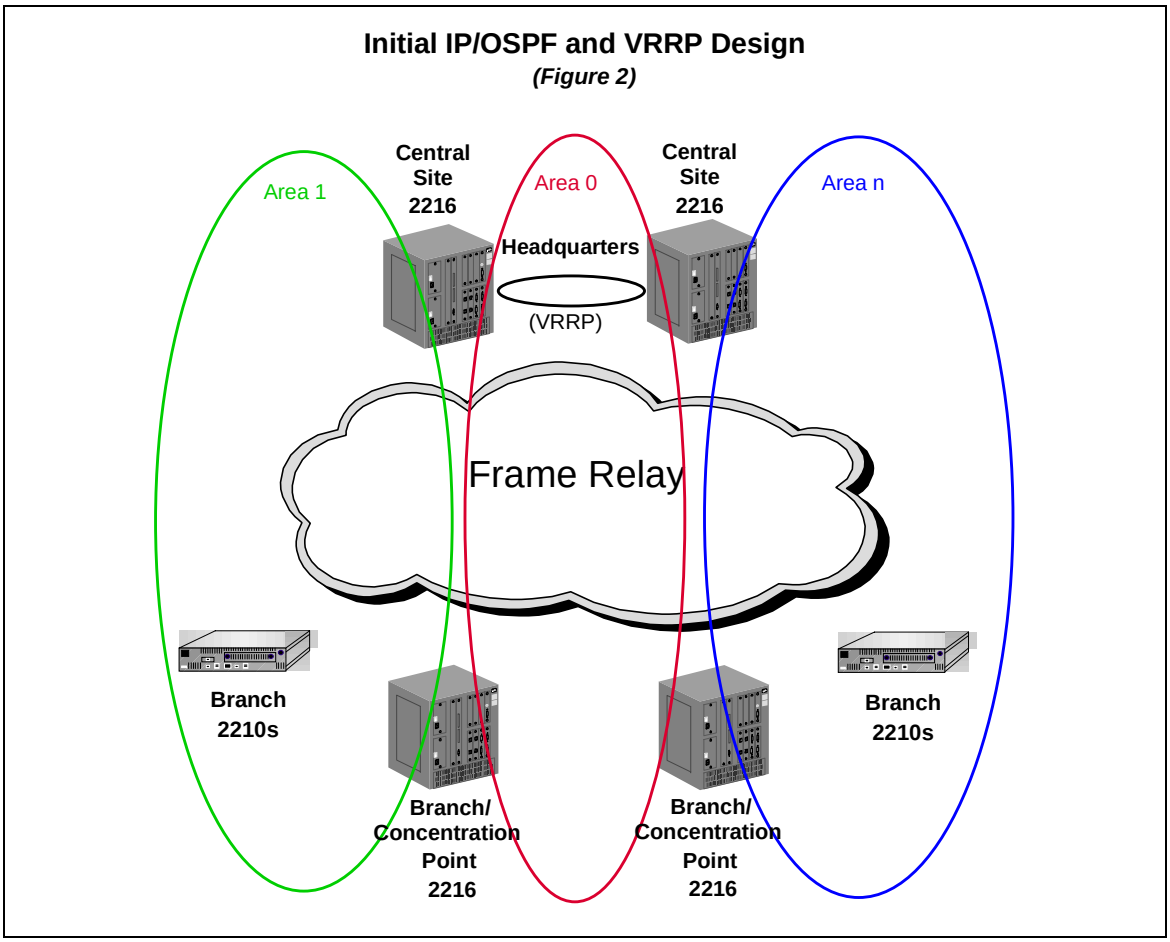
The new network will center around a pair of headquarter based 2216 M400 routers (hereafter also referred to as the *Central Site* routers) connected locally by a token ring segment. Also attached to that ring will be the FEPs/mainframes and servers accessed by remote site 2210 and 2216 routers. The remote sites will be organized into areas each containing 10 to 20 2210 Branch Office routers. The 2210s are each connected via Frame Relay PVCs (shown in the above figure by the colored lines) to:

- a Central Site router
- another 2216 router which acts both as a Concentration Point for the area's 2210s and as a Branch Office router for its own location.

Each Concentration Point router will be connected to the central site via Frame Relay PVCs to each of the Central Site routers. There will be about 20 concentration point routers in the network, one per area.

With this configuration, each router will have two PVCs providing, via OSPF, alternate paths in failure scenarios.

Initial IP/OSPF and VRRP Design



In order to achieve the required availability, the fast converging OSPF routing protocol was the obvious choice. To limit the size of routing tables and OSPF flows, a network of this size requires multiple areas and address summarization. Per the above logical diagram:

- Each 2216 will be a Backbone Border Router (BBR)
  - Each Central Site router will be in both Area 0.0.0.0 and half of the other areas in the network.
  - Concentration Point routers be in both Area 0.0.0.0 and their own areas.
- Each Branch Office 2210 will be an Area Router attaching to one of the Central Site Routers as well as its area's Concentration Point router.

Area addresses will be summarized in the backbone.

The servers attached to the Central Site token ring will specify one of the Central Site Router's token ring port addresses as their Default Gateway (DFG). To address the scenario in which that

port (or its router) fails, Virtual Router Redundancy Protocol (VRRP) will be configured for the token ring ports on the Central Site Routers.

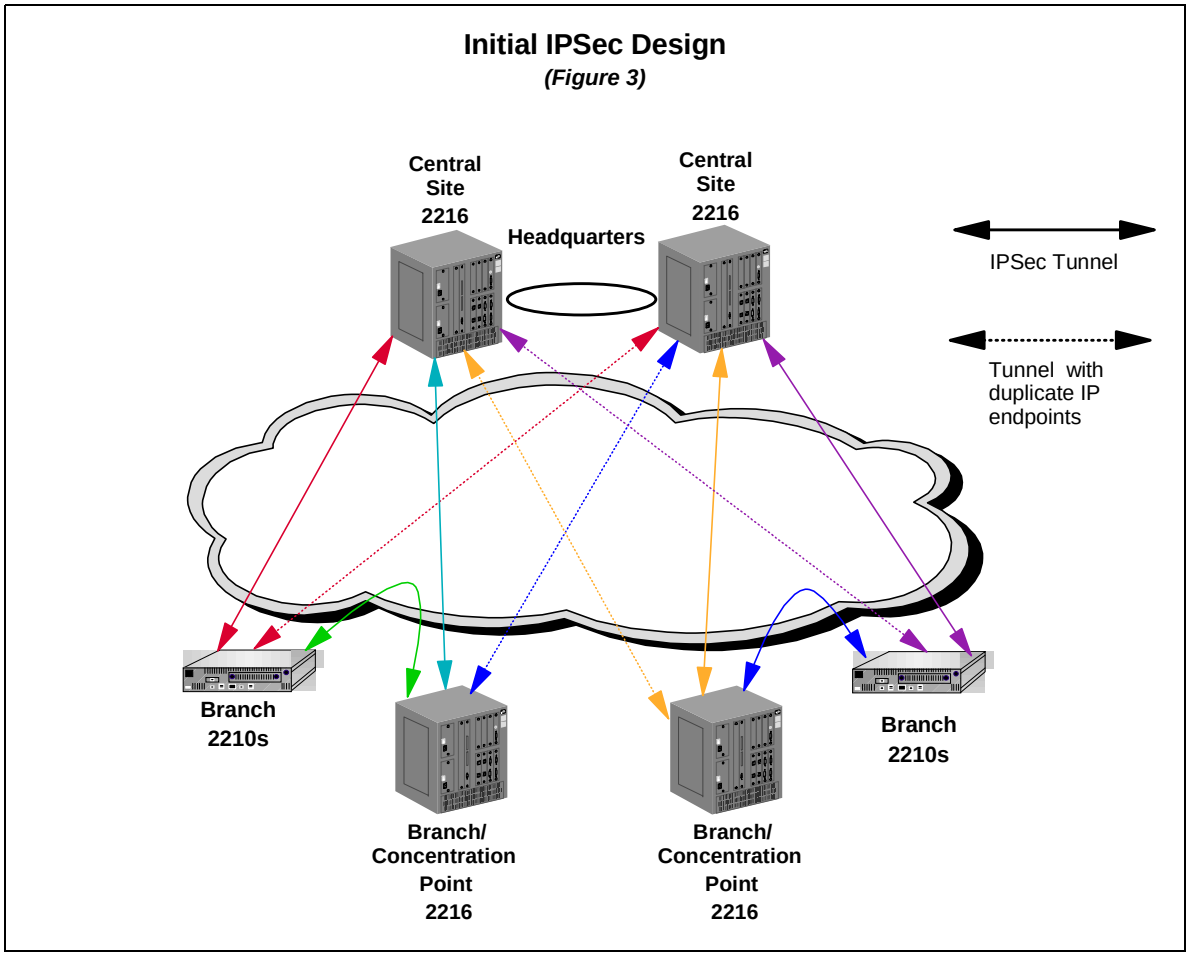
As will be seen later (and as many readers have probably already guessed), problems were encountered with this part of the design.

Initial IPSec Design

Privacy and integrity were necessary for data flowing between clients and servers located in branch offices as well as between clients in branch offices and the centrally located mainframes and servers.

Due to the use of PVCs, Frame Relay networks avoid many of the exposures for misdirected frames that can occur in a router network built on a PPP infrastructure. Unfortunately, the danger of an unauthorized third party viewing and possibly modifying sensitive data are as much an exposure in a public Frame Relay network as it is in a PPP based router network. There are several different approaches for addressing these threats.

- Link Encryption: Though this could provide the necessary security, it's a point to point technology intended for lease lines and would not be appropriate for a network with multiple PVCs/paths per physical interface.
- End Point Security: Security could be implemented on the end point devices (clients and servers) but this would require software upgrades on all these devices.
- IPSec: Since all backbone traffic was to be carried via IP, IPSec was a natural choice for meeting the customer's requirements. It was a better choice than Line Encryption since it operates at Layer 3 and can provide security regardless of which Frame Relay PVC/path actually carries the traffic. It was is also able to provide the security without requiring software upgrades on the clients and servers.



As shown in *Figure 3*, three types of IPSec Tunnels will be defined:

- between each Branch Office 2210 and a Central Site 2216
- between each Branch Office 2210 and its associated Concentration Point 2216
- between each Concentration Point 2216 and each Central Site 2216.

In order to provide the same level of resiliency for IPSec as was provided by the Frame Relay and OSPF infrastructures, duplicate IP addresses for tunnel end points were to be configured on the Central Site 2216s. These logically "duplicate" tunnels are indicated by the dotted lines. An in-depth explanation of the need for this configuration and how it was implemented is provided later in the document.

Setup, Implementation, and Testing

Test Overview

A small test network was constructed to validate key design elements and provide implementation recommendations. The objectives for each of the main design points were:

- **IP/OSPF**: Validate the scheme for multiple areas, address summarization, any to any routing, and rerouting around Frame Relay PVC and Central Site hardware failures.
- **IPSec**
  - Security: Ensure that privacy and integrity were provided for data flowing between clients and servers located in branch offices as well as between clients in branch offices and the centrally located mainframes and servers.
  - Efficiency: Develop a scheme to minimize IPSec related CPU utilization.
  - Ease of Configuration: Develop an access control template which would both standardize configurations as much as possible and minimize the amount of configuration necessary when new routers were added to the network.
- **Availability**:
  - Verify delivery of data in spite of:
    - complete loss of either Central Site router
    - loss of a Central Site router's token ring or WAN ports
    - loss of one of a Branch or Concentration Point router's two Frame Relay PVCs
  - Verify that the implementation of VRRP on the Central Site routers would provide backup Default Gateway functionality for devices attached to the Central Site token ring.
  - Confirm that the use of duplicate IPSec tunnel end point addresses will provide fault tolerant IPSec capability.

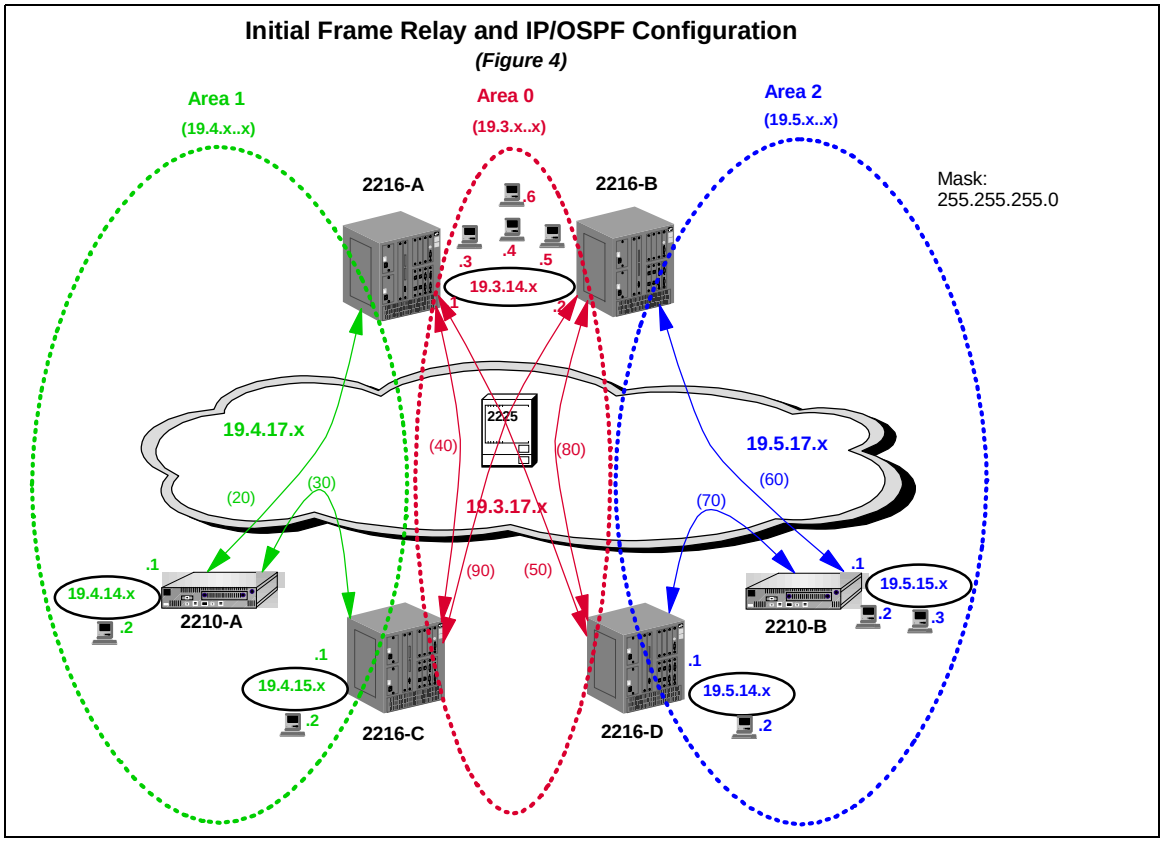
The functions involved allowed a staged implementation and test which greatly simplified the project. The system was implemented and tested in the following order:

1. Frame Relay and IP/OSPF infrastructure
2. VRRP
3. IPSec

Detailed test procedures were written for each phase of the testing. A copy of the final set of test procedures (testing the system with all functionality added) can be found in *Appendix A* (page 51).

Frame Relay and IP/OSPF Infrastructure Implementation/Testing

Initial Frame Relay and IP/OSPF Configurations



The basic test system consisted of four 2216 M400s and two 2210 M24Ts. Two of the 2216s (A and B) represented the Central Site Routers, while the other two (C and D) were Concentration Point routers. The two 2210s (A and B) acted as Branch Office routers. Branch and Concentration Point routers were connected to each other and the Central Site routers via Frame Relay PVCs. The two Central Site routers were token ring connected.

As implied by *Figure 4*'s title, the proposed design did not meet some of the key requirements mentioned earlier. As there is benefit in seeing what didn't work (and why) as well as what did work, both the initial unsuccessful configurations as well as the final ones will be discussed. Prior to going into the testing results, a more detailed explanation of the configuration of the system components will be provided.

Frame Relay Configuration

Frame Relay Configuration Overview

In *Figure 4* (page 14) the paths of the PVCs are indicated by the double arrow lines. The circuit numbers (really the DLCI numbers, the same being used at both ends of the circuit) are noted in parentheses next to the lines.

The Frame Relay configuration was designed to provide both high availability and efficiency.

- Each Branch Office router has two PVCs. One provides a path to its Concentration Point router and the other to its Central Site router.
- Each Concentration Point Router has two PVC, one to each of the Central Site routers.

The availability requirement for the Frame Relay backbone was that there would be no single point of failure which would inhibit Branch Office and Concentration Point routers from accessing each other or the centrally located mainframes and hosts. The implemented scheme maintains complete Frame Relay connectivity between all the routers (and the mainframes, servers, and hosts behind them) in spite of the loss of one PVC in a Branch Office and/or its Concentration Point router. Though failure of a Central Site WAN Port (or even the entire router) is covered, failure of a Branch Office or Concentration Point router WAN port (or connection to the service provider's Frame Relay switch) is not. This failure scenario will be addressed by some form of analog or digital dial out capability which, as mentioned earlier, was not included in this testing.

With regard to efficiency, the design minimized the number of PVCs that had to be supported by the Central Site routers while still providing alternate paths in case of failure. Branch Office routers have only one PVC to the Central Site but they could rely on the PVC to the Concentration Point Router for backup. Since there would be a significant amount of traffic going between Branch and Concentration Point routers, the PVCs configured directly between them (30 and 70) off loaded traffic and processing from the Central Site routers.

Frame Relay Configuration Details

The Frame Relay backbone was provided by a single IBM 2225 M400 Frame Relay switch. All the PVCs were configured port to port with the midplane acting as the Frame Relay "cloud". All switch V.35 interfaces were configured as DCEs, provided a line rate of 2.048mbps to the attached routers, and acted as the Network side for LMI Annex D. Since mainly ping traffic was planned the same simple traffic descriptors were used for all PVCs: CIR: 256kbps, BC: 256kb, BE: 0.

Each router had a single V.35 connection to the switch. In the actual network, only the Branch Office routers would actually have single physical connections as the bandwidth requirements for the Connection Point and Central Site routers would probably exceed the line rate of one interface.

Following are some of the Frame Relay configuration details for the routers:

- Physical Interfaces: DTE (external clocking)
- LMI: Annex D
- PVC Traffic Parameters:
  - PVC Traffic Descriptors: CIR: 256kbps, BC: 256kb, BE: 0.
  - Traffic Control: All PVCs used the default configuration value of "Congestion Monitor". Due to the limited traffic volumes no congestion or frame discards were expected or encountered.

Frame Relay Configuration Results

The Frame Relay design met all its requirements and was maintained in its initial form for the duration of the testing. However, PVC redundancy is of little value in an IP network unless the routing protocol can correctly reroute the data flows to take advantage of those circuits. The proposed IP/OSPF design will be discussed next.

IP Configuration

IP Configuration Overview

The IP addressing scheme was chosen to facilitate summarization and testing (rather than for efficient use of the IP address space). Area addresses were identified by the second octet: Area 0.0.0.0 is 19.3.0.0, Area 0.0.0.1 is 19.4.0.0, and Area 0.0.0.2 is 19.5.0.0. WAN interface addresses were identified by a ".17" in the third octet while LAN addresses were identified by a ".14" or ".15" in the third octet. Hosts addresses on each subnet were identified by the fourth octet. The subnet mask was 255.255.255.0.

IP Configuration Details

There are only a couple of configuration details that need to be mentioned:

- ***inARP*** was used to automatically map Frame Relay PVCs to their IP addresses. This was done to eliminate the need for manual configuration of the mapping as routers were added to the network.
- The addresses used for each routers' Internal Address had an importance IPSec implication which will be discussed later (*Tunnel Configuration Details*, page 39).

IP Configuration Results

The addressing scheme described above proved successful and was not modified during the testing.

OSPF Configuration

Consistent with the overall network objectives, the main goals for the OSPF implementation were efficiency and availability.

OSPF efficiency involves several different elements. The use of the lowest cost paths as well as load balancing over equal cost paths are both provided as part of basic OSPF functionality. Minimizing OSPF related network and router overhead, however, is dependent on network design and configuration. Limiting the size of routing tables and amount of topology maintenance traffic is achieved though the use of multiple areas and address summarization. The goal is to limit an area's knowledge of other areas to the minimum amount of information needed for the hosts to communicate.

With regard to availability, OSPF is well known for its quick convergence time and ability to reroute around failures with minimal disruption to the traffic. Unfortunately, as will be shown later, the objectives of efficiency and availability sometimes conflict because address summarization often results in the use of sub-optimal routes as well as providing less robust rerouting capabilities.

OSPF Configuration Overview

The test network was divided into three OSPF areas as indicated in *Figure 4* (page 14) by color coded markings.

- Area 0.0.0.0 (Red): This is the 19.3.0.0 network which includes each 2216 as a Backbone Border Router (BBR). All the IP connections between the routers in this area ran over the 19.3.14.0 token ring and the 19.3.17.0 Frame Relay network (PVCs 40, 50, 80, and 90). Four NT workstations were attached to the 19.3.14.0 token ring to act as traffic sources and sinks. All addresses were summarized in the BBRs at the area level (19.3.0.0, 19.4.0.0, 19.5.0.0).
- Area 0.0.0.1 (Green): This is the 19.4.0.0 network which includes 2216-A, 2216-C , and 2210-A. The IP connections between these routers in Area 0.0.0.1 ran over the 19.4.17.0 Frame Relay network (PVCs 20 and 30). 2210-A and 2216-C each had an attached token ring (19.4.14.0 and 19.4.15.0) with one NT workstion to act as a traffic source and sink.
- Area 0.0.0.2 (Blue): This is the 19.5.0.0 network which includes 2216-B, 2216-D, and 2210-B. The IP connections between these routers in Area 0.0.0.2 ran over the 19.5.17.0 Frame Relay network (PVCs 60 and 70). 2210-B and 2216-D each had an attached token ring (19.5.15.0 and 19.5.14.0). 2216-D had one NT workstation while 2210-B had two to act as traffic sources and sinks.

OSPF Configuration Details

Following are some of the key OSPF configuration details:

- Addresses were summarized in each BBR based on the first two octets of the address (19.3.0.0, 19.4.0.0, and 19.5.0.0).

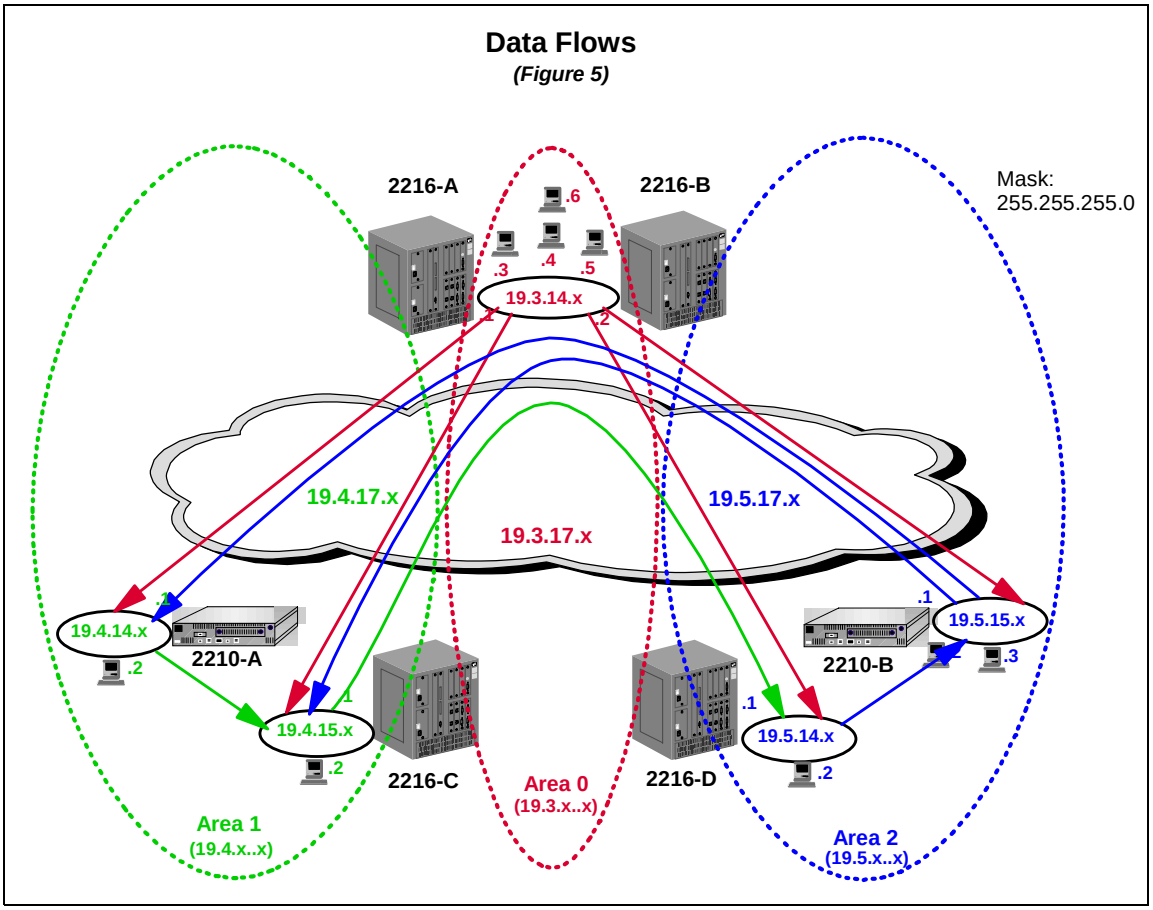
- Configuration of Designated Router priorities were not necessary. Since the Frame Relay network is partially meshed, the “Point-to-Multipoint” default was used.
- OSPF Neighbors only need to be identified at one end of a connection so to minimize configuration activity as new routers were added to the network:
  - Neighbors were not configured on the Central Site Routers.
  - Concentration Point Routers identified only the Central Site Routers as neighbors.
  - Branch Office Routers identified the Central Site router in its area and its Concentration Point Router as neighbors.
- The default values were used for all OSPF keep alive timers and counters.
- The OSPF interface *Costs* were assigned based on media type, Token Ring or Serial, and speed. Since all Frame Relay PVCs had the same AccessRate/CIR all had equal costs. Predicting flow patterns in a multiple area addressed summarized network can be difficult as forward paths are often different than the return. This situation will be further complicated when host site Default Gateways and IPSec Tunnel End Points are considered since these will tend to force traffic in specific directions. A detailed traffic and flow analysis (not part of this test) will be needed to accurately provision bandwidth for the Frame Relay PVC and make informed OSPF Load Balancing decisions.

OSPF Configuration Testing Results

As mentioned earlier, the Frame Relay and IP configuration were not changed for the duration of test. This was not the case for the OSPF design. As will be discussed next, the initially proposed design did **not** meet the basic network objectives and significant modifications were required.

Test Procedure Overview

Prior to reviewing the results of the initial testing, a brief overview of the manner in which the system was tested is in order.



Since the purpose of the testing was only to validate functionality, traffic was kept simple, consisting of pings, telnets, and the uploading of router configurations through the network via SNMP. As mentioned earlier, traffic flow was supposed to continue with minimal loss in spite of the failure of any one Frame Relay PVC or the complete loss of one of the Central Site routers, its WAN interface, or its token ring interface. Testing of the rerouting of traffic to/from Central Site work stations after token ring adapter failure was not done until after VRRP was implemented.

Figure 5 shows the nine flows of the pings that were used in the testing (sorry for the busy diagram). Each work station acted as data source and sink for the others. The lines indicate both the source and destination of the pings (the arrow heads point toward the destinations) while the colors indicate the OSPF area **from** which the flow originated. The flows represented traffic from the Central Site to Areas 0.0.0.1 and 0.0.0.2, between Areas 0.0.0.1 and 0.0.0.2, and within Areas 0.0.0.1 and 0.0.0.2.

The system was tested in the following stages with appropriate configuration changes made after each:

1. steady state without address summarization
2. failure scenarios without address summarization
3. steady state with address summarization
4. failure scenarios with address summarization

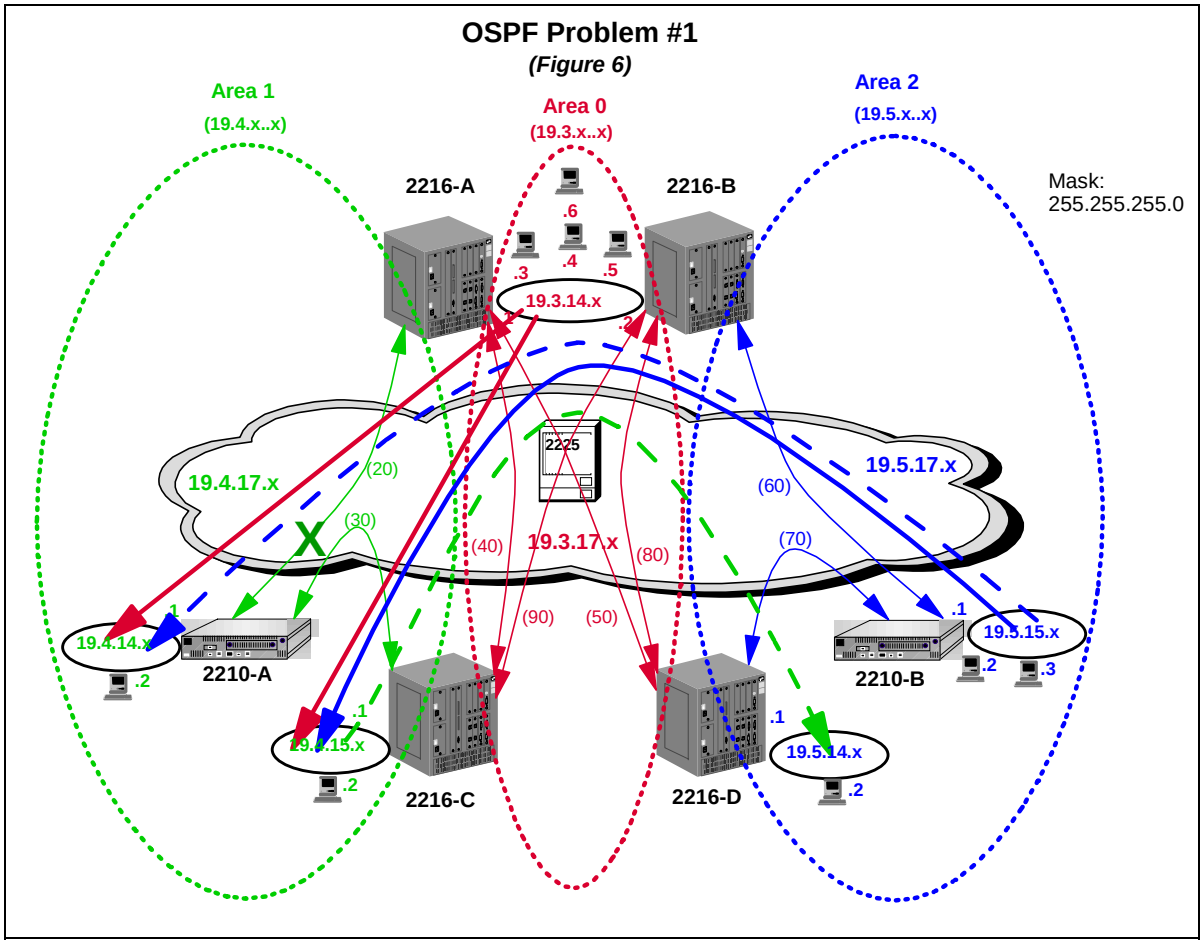
The results of the testing are next.

Initial IP/OSPF Configuration Test Results

Everything worked perfectly without address summarization in both steady state and failure scenarios. Interrupted traffic continued within about 40 seconds after the disabling of any PVC (from the Frame Relay switch's NMS console) or removing the WAN cables from either of the Central Site routers.

The address summarization described earlier was then added to the routers' configurations and the tests rerun. Steady state again worked perfectly. Unfortunately, rerouting problems were encountered as soon as the first PVC was disabled.

OSPF Rerouting Problem



As soon as the first Frame Relay PVC (20) was disabled five of the nine traffic flows (shown above) stopped. There were two ping failure reasons which appeared on the monitors of affected work stations:

- *Time Out*: This occurred for the ping flow, shown by the segmented green line, from 2216-C's workstation (19.4.15.2, Area 0.0.0.1) to 2216-D's workstation (19.5.14.2, Area

0.0.0.2). Remember, PVC 20 had been disabled. To get to the target the pings would use the following Frame Relay PVCs: 90, 60, 70. Even though there was a shorter path (90, 80), per standard OSPF operation 2216-B will always select a path that is in the area of the target if one is available even if it is not the most efficient. PVC 80 is in Area 0.0.0.0 so PVC 60 is 2216-B's only circuit in Area 0.0.0.2. So far so good, but the response to the ping still has to get back to the originator. Since the WAN port on 2216-A is still active, due to address summarization the router continues to advertise reachability to "anything" in the 19.4.0.0 network. Since PVC 50 represents the shortest path to a BBR that has an interface in the target area (0.0.0.1), 2216-D sends the ping response to 2216-A via that route. 2216-A will then tell 2216-D "*I know I told you I can get to destinations in 19.4.0.0 but unfortunately the specific destination, 19.4.15.2, you want to send the ping response to is unreachable*". Accordingly, 2216-A returns a *Destination Unreachable* response to 2216-D and the *Time Out* message appears on the monitor of the work station that originated the ping since it didn't get a response within the ping's specified time frame.

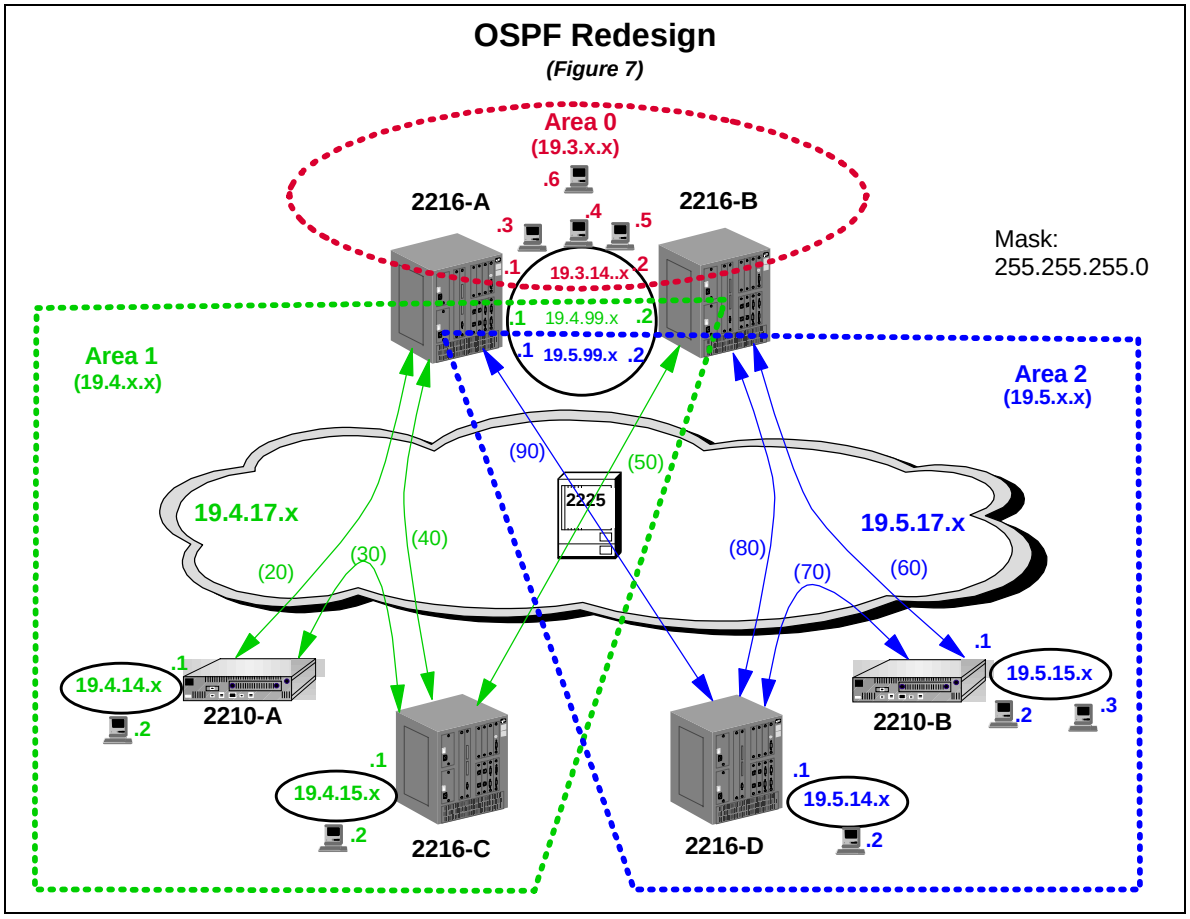
- *Destination Unreachable*: This occurred on four of the failing ping flows; describing one should be sufficient. One of 2210-B's work stations (19.5.15.2, Area 0.0.0.2) was pinging 2210-A's workstation (19.4.14.2, Area 0.0.0.1) as shown by the segmented blue line. Prior to disabling PVC 20, 2210-B used Frame Relay PVCs 70, 50, 20 to reach the destination. Once PVC 20 was disabled we were in exactly the same situation as the *Time Out* problem described above. Since the problem occurred on the outbound path of the ping, vs. the response, a *Destination Unreachable* response was returned to the source workstation.

Similar problems occurred after disabling other PVCs. The initial OSPF design was obviously unacceptable. The changes made to correct the problems are explained next.

OSPF Configuration Redesign

To get around the problems with the initial configuration and still use address summarization each BBR would need at least two PVCs into an Area so there would be an alternate path in case one failed. The initial design could have been adjusted but it would have meant supporting Area 0.0.0.1 and Area 0.0.0.2 IP addresses on the PVCs that formerly only had Area 0.0.0.0 addresses (PVCs 40, 50, 80, and 90, see *Figure 4* on page 14). Unfortunately, since this would mean some PVCs would be supporting multiple subnets, ***inARP*** could ***not*** be used because only the first "discovered" address would be cached. The addresses for each area (and there would be about 20 in the actual customer network) would have to be added manually at each end of every PVC in the Area 0.0.0.0. backbone. The effort needed to make the original design meet the network requirements caused a reassessment of the overall OSPF design.

OSPF Redesign Overview



Due to the amount of manual configuration needed to provide the required rerouting capability and still keep every Concentration Point router a Backbone Border Router, it was decided to modify the design. The network was still divided into three OSPF areas as indicated in the *Figure*

7 above but now only the Central Site Routers were Backbone Border Routers. The Concentration Point Routers were now, like the 2210s, just Area Routers.

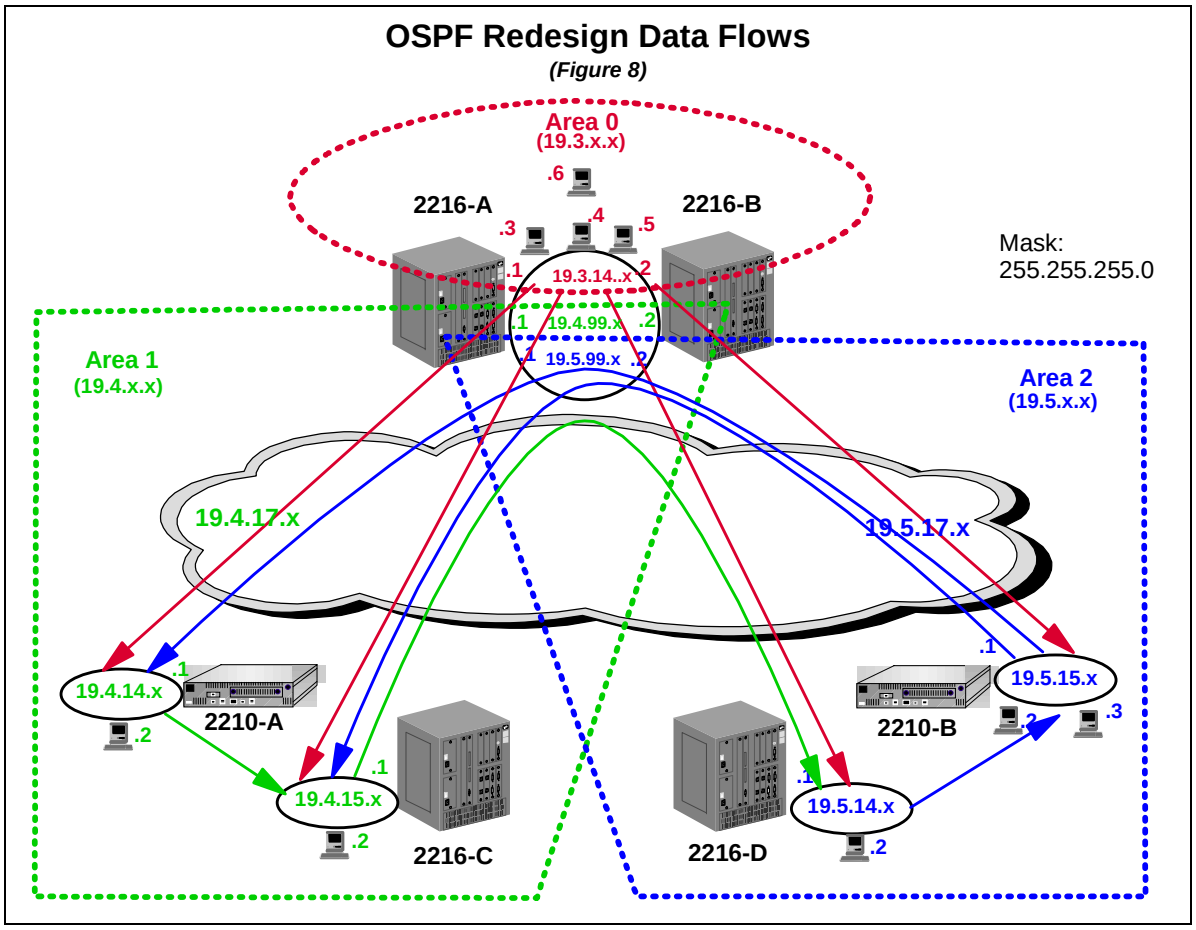
- Area 0.0.0.0 (Red): This is the 19.3.0.0 network which includes 2216-A and 2216-B as Backbone Border Routers. Area 0.0.0.1 and Area 0.0.0.2 addresses (19.4.99.1/.2 and 19.5.99.1/.2) were added to these routers' token ring interfaces to allow rerouting across the ring. To keep rerouted traffic off the same ring which provides access to the Central Site servers and mainframes, a separate ring between the two routers could have been added just for rerouting purposes though this was not done for this test. There are no longer any Frame Relay PVCs in Area 0.0.0.0. Only Area 0.0.0.1 and Area 0.0.0.2 addresses were summarized (as before on the first two address octets). Area 0.0.0.0 addresses were not summarized to avoid scenarios in which one valid Area 0.0.0.0 address (i.e. possibly the Internal Address) could cause the router to advertise reachability to the central ring's subnet when those resources were really unreachable. Though this results in the appearance of all Area 0.0.0.0 addresses in the other areas' routers' routing tables this was not considered a major performance impact as the total number of those entries would be small.
- Area 0.0.0.1 (Green): This is the 19.4.0.0 network which includes 2216-A, 2216-B, 2216-C and 2210-A. The IP connections between these routers are all in Area 0.0.0.1 running over the 19.4.17.0 Frame Relay network (PVCs 20, 30, 40, and 50). The main difference between this and the previous configuration of the area is that the two Frame Relay PVCs connecting 2216-C to the Central Site routers (40 and 50) are now in Area 0.0.0.1 instead of Area 0.0.0.0.
- Area 0.0.0.2 (Blue): This is the 19.5.0.0 network which includes 2216-A, 2216-B, 2216-D, and 2210-B. The IP connections between these routers in area 0.0.0.2 ran over the 19.5.17.0 Frame Relay network (PVCs 60, 70, 80, and 90). The main difference between this and the previous configuration of the area is that the two Frame Relay PVCs connecting 2216-D to the Central Site routers (80 and 90) are now in Area 0.0.0.2 instead of Area 0.0.0.0..

Note that the original Frame Relay configuration did not change. The PVCs between the 2216s are now just in different OSPF areas. Also, since each Frame Relay PVC is in only one subnet, ***inARP*** can still be used.

Other details of the new configuration are basically the same as the initial design's.

OSPF Redesign Test Results

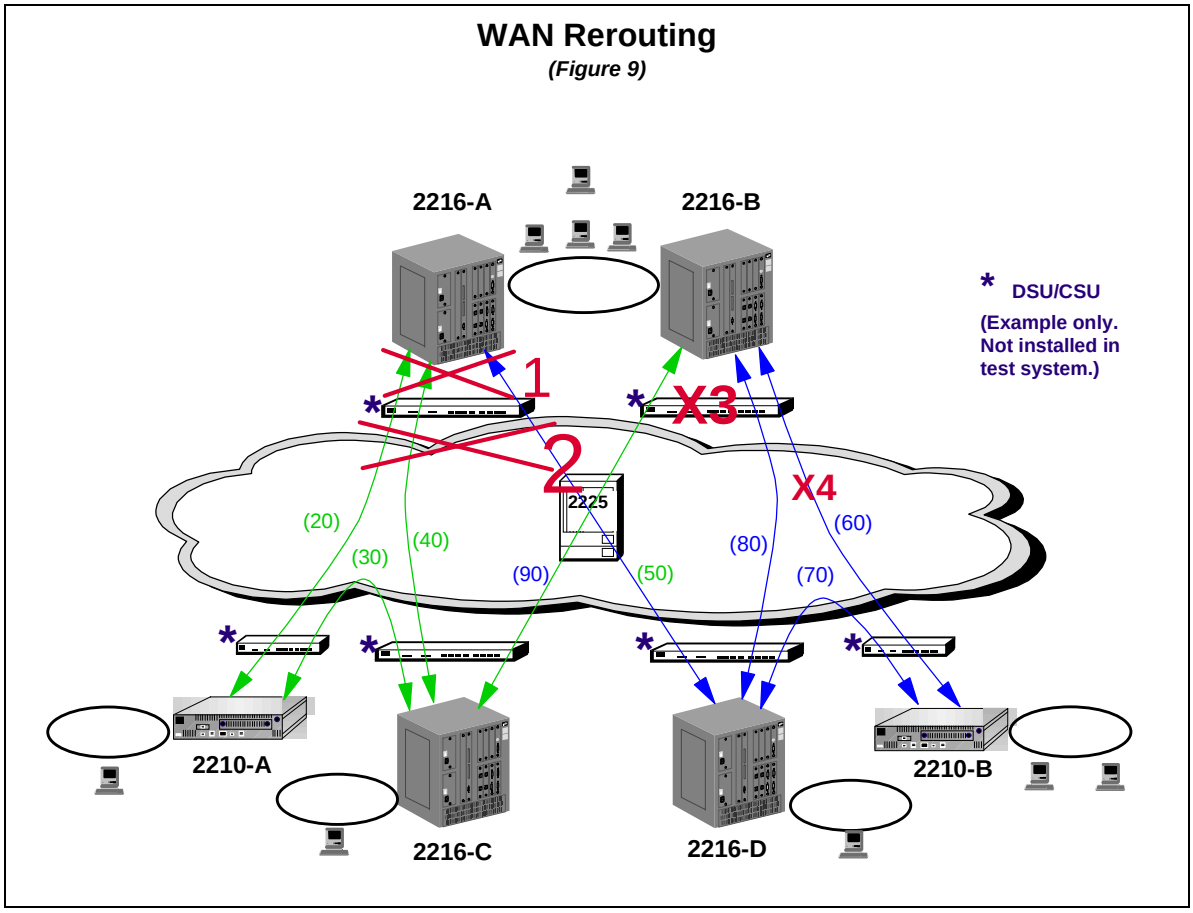
The following diagram shows the traffic flows (same as before) overlaid on the new configuration.



To validate the new configuration the test procedures were first run after changes were made to the basic OSPF Area configuration but prior to the implementation of address summarization. As had been the case with the initial non-summarized OSPF configuration all traffic flowed as desired in both the steady state and failure scenarios. Steady state operation was also successful after summarization was added. Unlike the initial OSPF configuration though, rerouting after PVC and physical interface failures **was** successful. As will be explained next, the length of time it took for disrupted traffic to resume was dependent on the type of disruption.

WAN Rerouting

Following is an overview of the different WAN failure scenarios that were tested.



The different scenarios are indicated by the numbered red *X*s. Note that DSU/CSUs are included in the diagram for explanation purposes (even though none were involved in the actual testing) since they would be installed in the customer's actual network.

- *X1*: This represents the failure of the router's WAN port which was simulated by removing the V.35 cable from the router. 2216-A would immediately inform the other routers in the network via OSPF (over token ring through 2216-B) that it no longer had direct access to Area 0.0.0.1 or Area 0.0.0.2. Pings that had been flowing through 2216-A's WAN port would only be interrupted for about 5 seconds before resuming.
- *X2*: This represents the failure of the Frame Relay Switch's WAN port. It could also represent the failure of the entire switch the router was connecting to at the service provider's point of presence. This produced the same results as the *X1* failure. Note that even though there would be a DSU/CSU between the router and Frame Relay switch disconnecting the cable (or port failure) at the router or switch causes the DSU/CSU to bring down its interface on the other side thus informing that device that the end to end physical interface was down.

- *X3*: This is the failure of a DSU/CSU. This would cause the involved router and Frame Relay switch to both realize the physical interface was down so the scenario is really the same as *X1* and *X2*.
- *X4*: This represents the failure of a single Frame Relay PVC. As mentioned earlier, this was accomplished by disabling the PVC at the Frame Relay switch's network management station. Since the Frame Relay switch and the PVC's port were still active when the PVC was disabled, the switch sends a "Circuit Down" LMI message to the routers at both ends of the PVC. Unlike the case when the router detects that the physical interface is down, the only way OSPF finds out that PVC 60 is down is when the OSPF "Dead Router" timer expires due to lack of Hello responses. There is a "Required Circuit" option for PVCs which tells the router to "bring down the interface if this PVC fails" but it would be inappropriate for this network since multiple PVCs are defined on the interface to provide alternate paths. Thus, reroute time will always be tied to the size of the "Dead Router" timer which for this test was the default value of 40 seconds. Reducing the timer's value to improve reroute performance would have to studied in the customer's environment as it could have undesirable effects on performance and PVC stability if it was made too small.

Testing of Central Site Router Token Ring interface failure and complete loss of a Central Site Router required the implementation of VRRP. This is described next.

VRRP Implementation/Testing

The workstations attached to the Central Site ring each specified the IP address of 2216-A's token ring port (19.3.14.1) as their Default Gateway (DFG) for traffic destined outside the 19.3.14.0 subnet. To allow the rerouting of traffic from those workstations in the event of the failure of 2216-A's token ring port (or the loss of the entire router) Virtual Router Redundancy Protocol (VRRP) was implemented on both Central Site Routers.

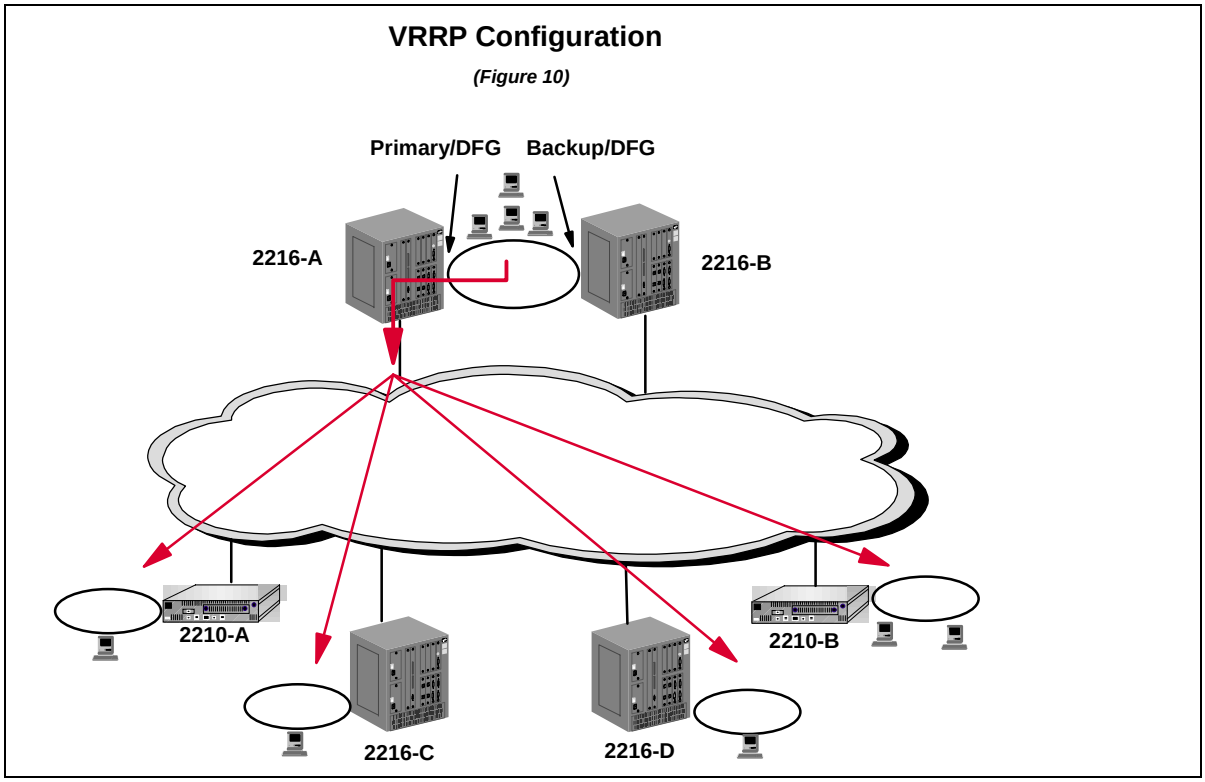
VRRP Configuration

2216-B's token ring port was configured to act as backup to 2216-A's token ring port. 2216-B will take over DFG responsibilities when it detects that 2216-A's token ring port is no longer operational. It decides 2216-A's port is down when it doesn't receive any "*Primary is Alive*" messages from that adapter within a certain time period (3 x the expected transmission interval). The configuration of the function is quite simple so the details will not be covered.

The following diagram shows the configuration just described as well as the *assumed* (reason for italics explained later) data flows from the Central Site ring attached work stations in a non-failure scenario.

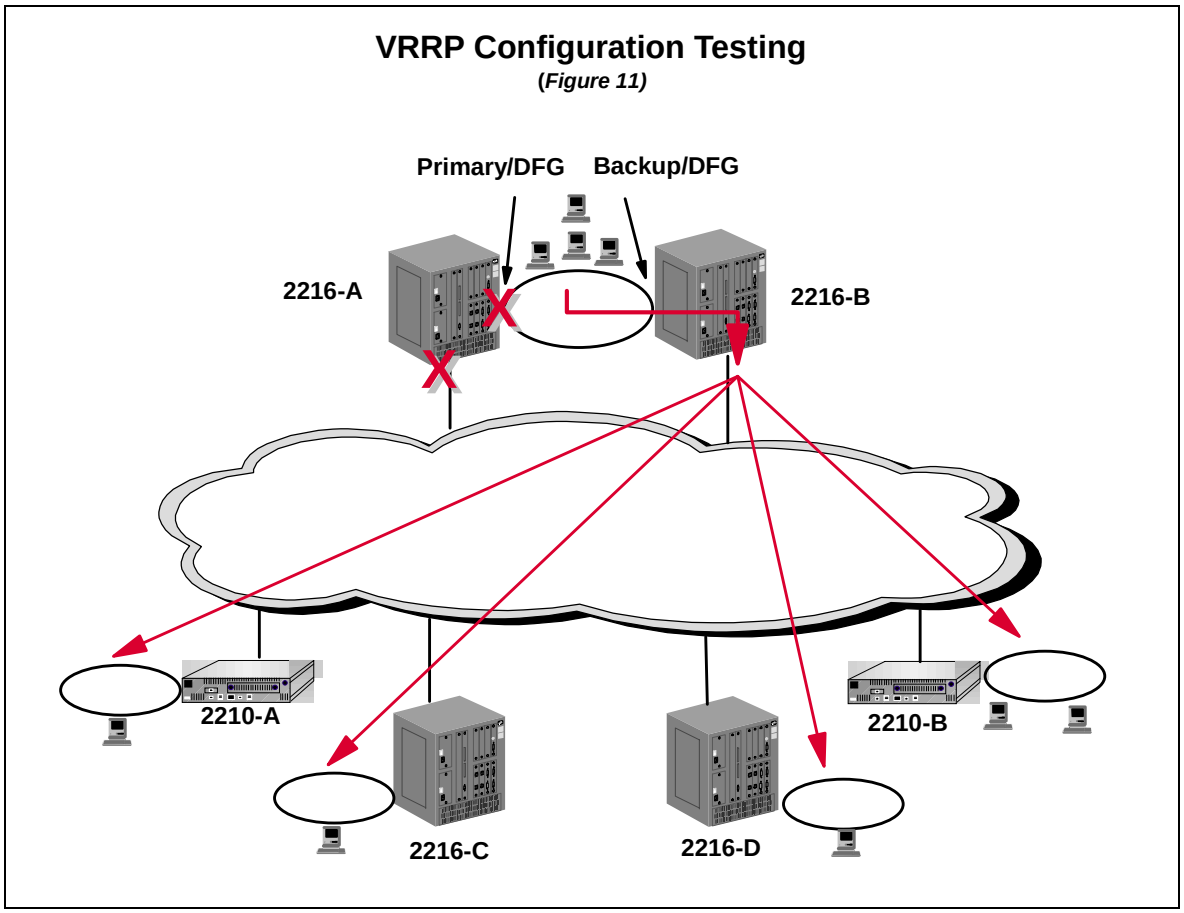
VRRP Configuration

(Figure 10)



VRRP Configuration Testing Results

The function was tested by initiating all the traffic flows used before (*Figure 8*, page 25) and simply disconnecting 2216-A's token ring cable (and later both the token ring cable and the WAN connection). As expected, the flows between Area 0.0.0.1 and Area 0.0.0.2 were unaffected. As hoped, per the following diagram, *all* the flows from the Central Site workstations continued after only a brief interruption of about 3 seconds.



Though the initial testing was promising, a couple of problems were encountered later.

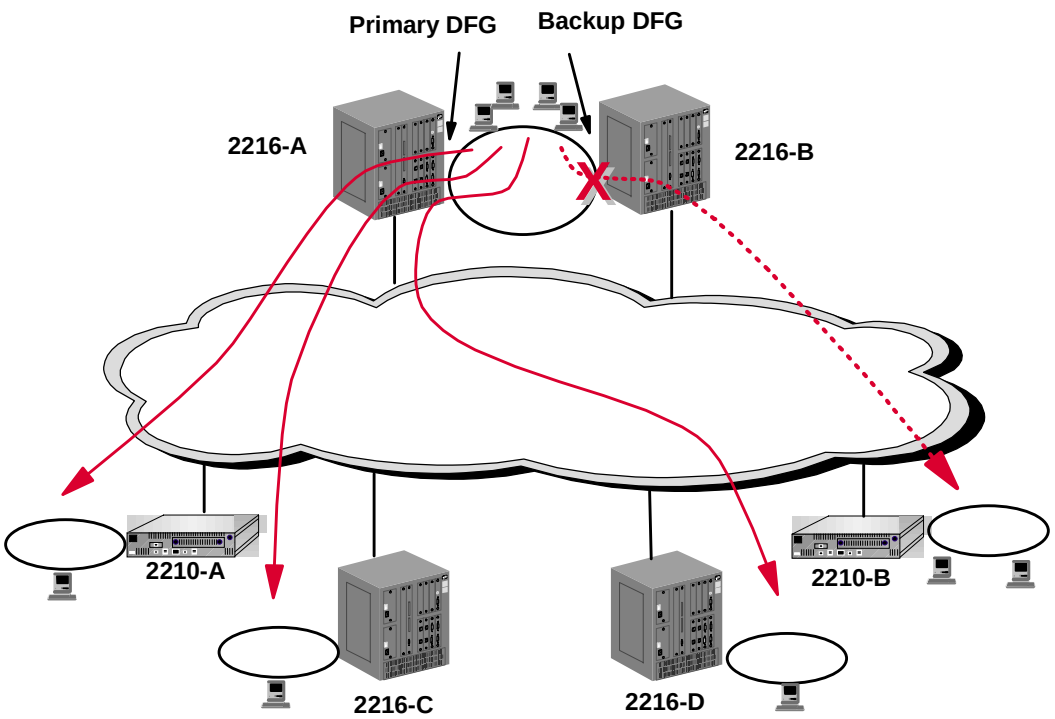
VRRP Problems/Solutions

VRRP Problem/Solution #1

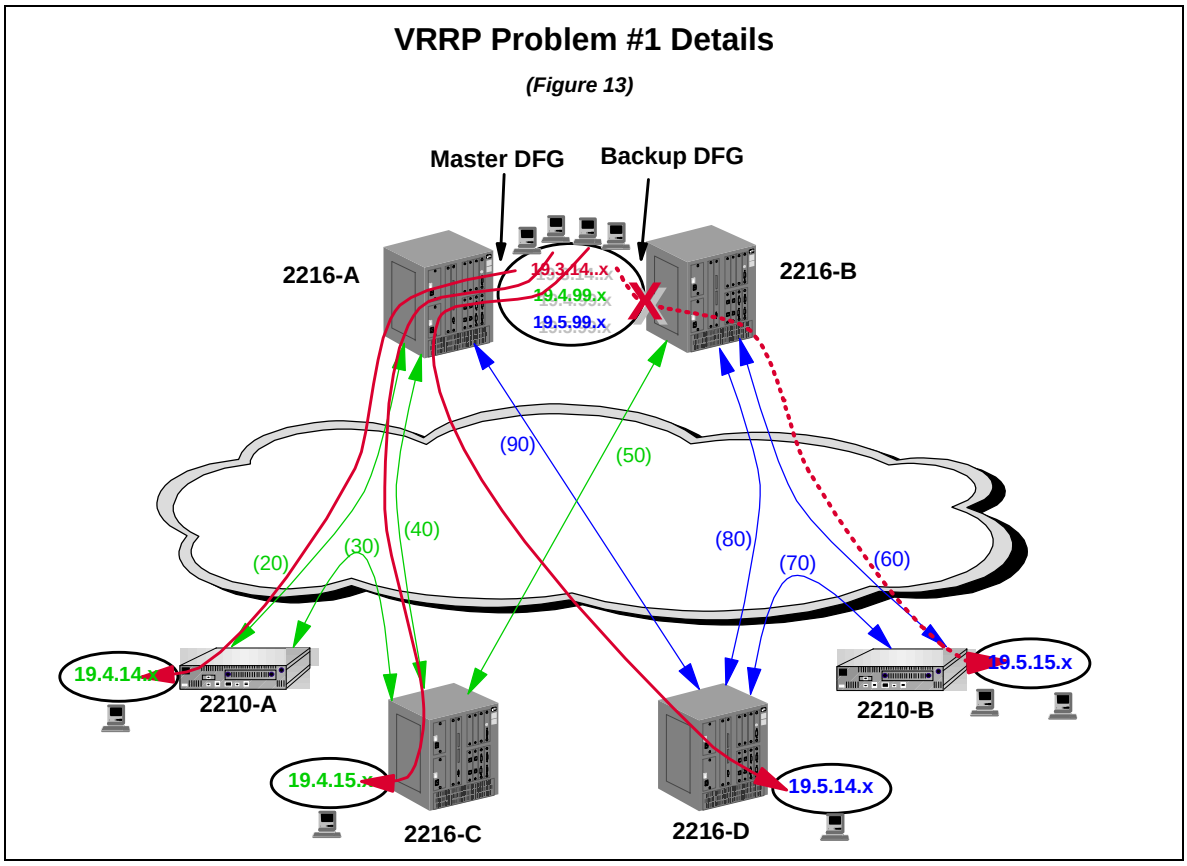
While testing the scenario of the failure of 2216-B's token ring port (really not part of VRRP testing) it was noticed that one of the data flows from a Central Site work station would take about 5 minutes to resume successful pinging to a workstation on 2210-B's token ring.

VRRP Problem #1

(Figure 12)



As shown in *Figure 12*, even though 2216-A's token ring port's IP address was the DFG for all the work stations on the Central Site ring, the pings to one of 2210-B's work stations (identified by the dotted line) were being sent to 2216-B's port thus causing the interruption of that traffic when that port was disabled. It's easier to understand what was happening if the problem scenario is overlaid on the Frame Relay components of the network as shown in Figure 13 on the next page.



First, let's look at the sequence of events prior to the disabling of 2216-B's token ring port:

- The ping to a workstation on the 19.5.15.0 subnet is issued by the Central Site workstation and is sent to the DFG on 2216-A.
- 2216-A determines that the shortest path to 2210-B's ring is via 2216-B (two interim WAN hops are more expensive than one token ring and one WAN hop) so the ping is sent back out onto the token ring to 2216-B.
- 2216-B forwards the ping to 2210-B via Frame Relay PVC 60.

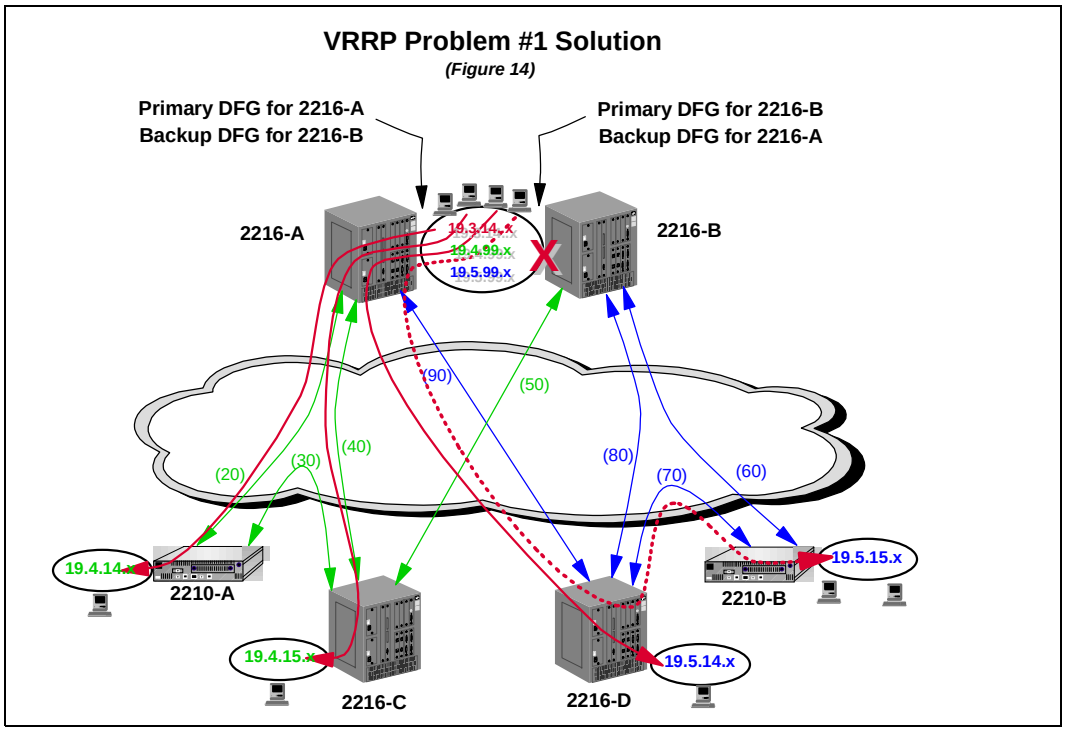
This explained why the traffic was traversing 2216-B but initial investigation into the cause of the problem raised a few additional questions:

- Why wasn't the traffic rerouted via 2216-A and 2216-D after the 40 seconds it should have taken OSPF in 2216-A to realize that 2216-B was no longer reachable over the ring?
    - The routing tables indicated 2216-A had realized 2216-B's token ring port had gone away and that the path over PVCs 90 and 70 was available.
    - All of the other workstations on the Central Site ring could ping the workstations on the 19.5.15.0 subnet.
- Why would the traffic restart as soon as 2216-B's token ring port was re-enabled or on its own (with the port still disabled) after about 5 minutes?

After much investigation it turned out the "problem" was the result of "ICMP Redirects". This is a per IP interface option (the default) which is implemented to avoid traffic from traversing the same LAN segment twice on the way to the ultimate destination (sound familiar?). In the situation under discussion, 2216-A realized that the workstation could go directly to 2216-B's token ring port to reach its destination so it sent a "redirect" message to that station to identify a new "preferred router" (DFG) for that destination. Thus, the system was working properly but one Central Site workstation had changed its default gateway to 2216-B's port which had only been configured to act as 2216-A's VRRP backup. The reason the traffic would start flowing on its own after about 5 minutes was due to the length of time it took for the workstation to age out its redirect instructions**.** If closer attention had been paid when the initial testing of VRRP had been done (see *Figure 10*, page 28) it would have been noticed that one of the Central Site's work station's pings had **not** been interrupted when 2216-A's token ring port had been disconnected. This was the work station that had been redirected to use 2216-B's token ring port so disconnecting 2216-A's port had no effect.

There were two possible ways to solve this problem:

1. Disable ICMP Redirects for the IP interface on the Central Site Token Ring.
2. Configure 2216-A's and 2216-B's token ring ports for VRRP so that in addition to 2216-B backing up 2216-A, 2216-A would back up 2216-B.



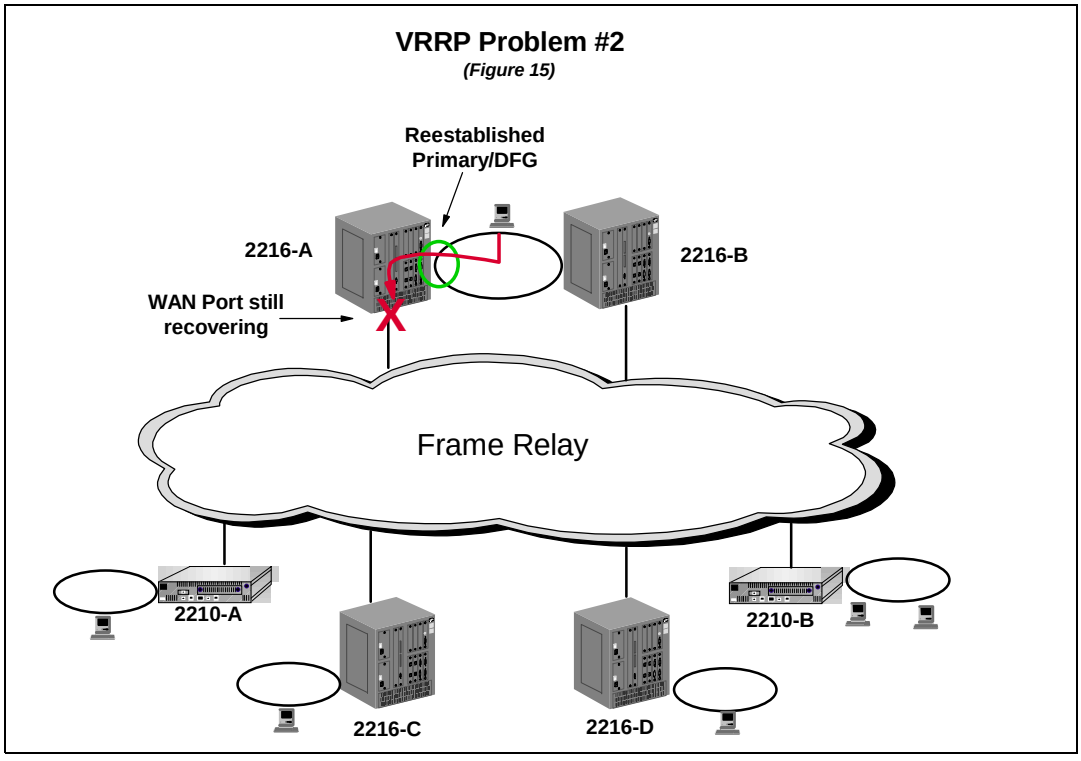
As show in *Figure 14*, the second option was chosen and testing confirmed it corrected the problem.

VRRP Problem/Solution #2

One of the key failure scenarios tested was the network's operation after the failure of one of the Central Site 2216s (simulated by simultaneously disconnecting its WAN **and** token ring cables). Though all the traffic rerouted correctly, whenever 2216-A was disconnected a slight problem was seen after that router was ***reconnected*** to the network.

Though no data should have been lost, pings from several of the workstations would be interrupted by about 4 failures with status indicating 2216-A felt the network for the IP address was *unreachabl*e. The pings would always resume after this brief interruption. After further investigation it was discovered this scenario would occur only if 2216-A's token ring port recovered ***before*** its V.35 port (and associated OSPF connections). If the V.35 port and OSPF connections recovered first no pings were ever lost.

Though I'm not 100% certain, I think what is pictured in *Figure 15* was occurring when the token ring port recovered before the WAN port.



The sequence of events would be as follows:

1. Disconnect 2216-A's WAN and Token Ring ports.
2. 2216-B's Token Ring port learns that 2216-A's port has gone down when it no longer receives VRRP "primary is alive" messages from that port. All traffic is rerouted around the failure by OSPF and VRRP. Now all traffic is flowing through 2216-B.
3. 2216-A's Token Ring and WAN ports are reconnected. The Token Ring port recovers first.

4. VRRP recovers 2216-A's "default gateway" responsibilities and the workstations' traffic starts being received by 2216-A. Unfortunately, 2216-A's WAN port is still initializing and the OSPF adjacency with 2216-B over the token ring has not yet been reestablished. During this window all pings sent to 2216-A from the central site work stations will result in "network unreachable" messages.
5. As soon as 2216-A's WAN interface (and associated Frame Relay PVCs and related OSPF relationships) recovers **or** the OSPF relationship with 2216-B reestablishes over the token ring, 2216-A can begin routing the pings.

Unfortunately there's really nothing that can be done to the routers' configurations to address this scenario but since the traffic is only briefly interrupted rather than terminated it's not a serious problem. The situation can be eliminated procedurally, when possible, by just bringing up the WAN port completely before reestablishing the token ring port.

IPSec Design, Implementation, and Testing

IPSec processing is closely tied to a router's Input and Output Access Control functions which are more commonly referred to as input and output filters. These filters are used to determine what traffic should be placed into a particular tunnel (output filters) and to validate which tunnel specific traffic should be received in (input filters).

Prior to delving into the configuration decisions made for the IPSec implementation, a brief explanation of how the input and output filters operate with regard to IPSec will help in understanding the rationale behind some of those decisions.

IPSec Output/Input Processing

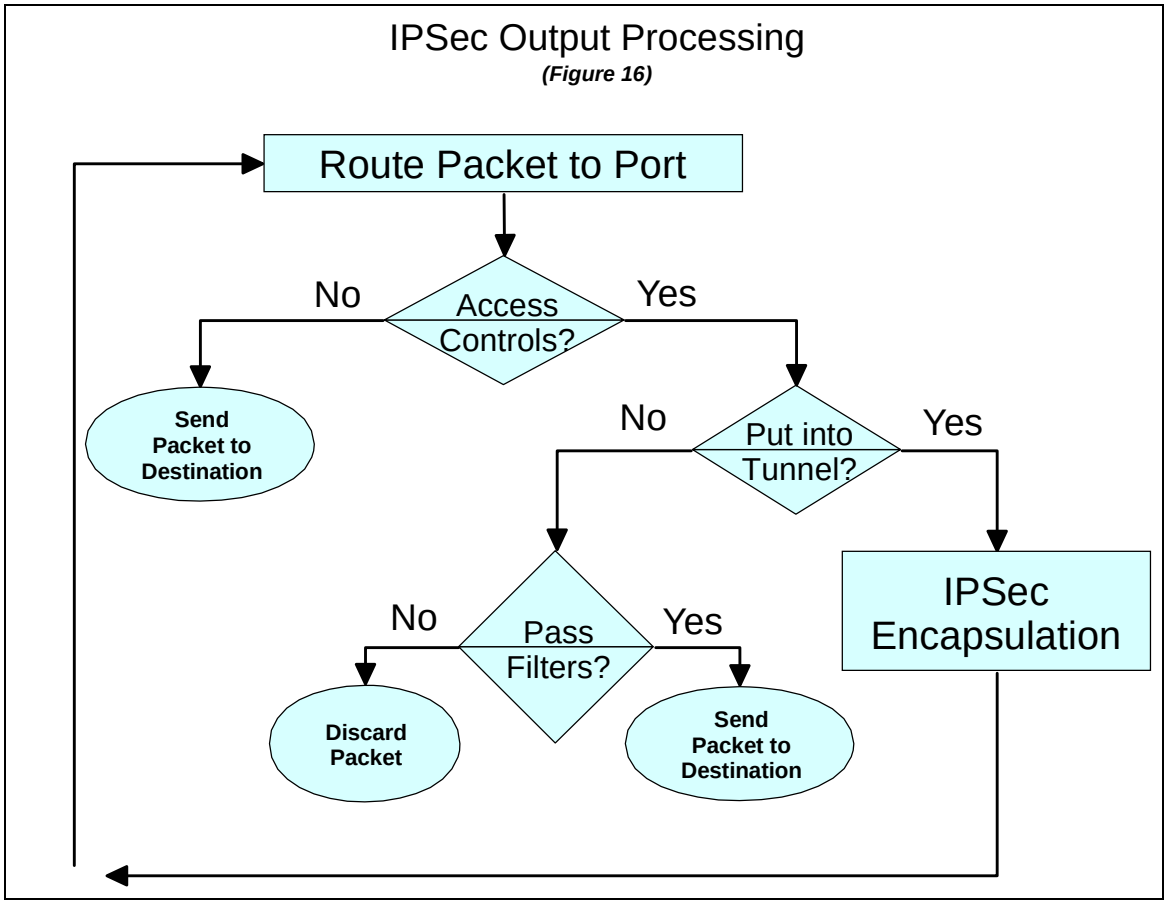
These explanations assume IPSec "Tunnel Mode" (vs. "Transport Mode") as Tunnel Mode was used in the implementation.

IPSec Output Processing

The following diagram shows the basic flow of port output processing for IPSec.

IPSec Output Processing

(Figure 16)

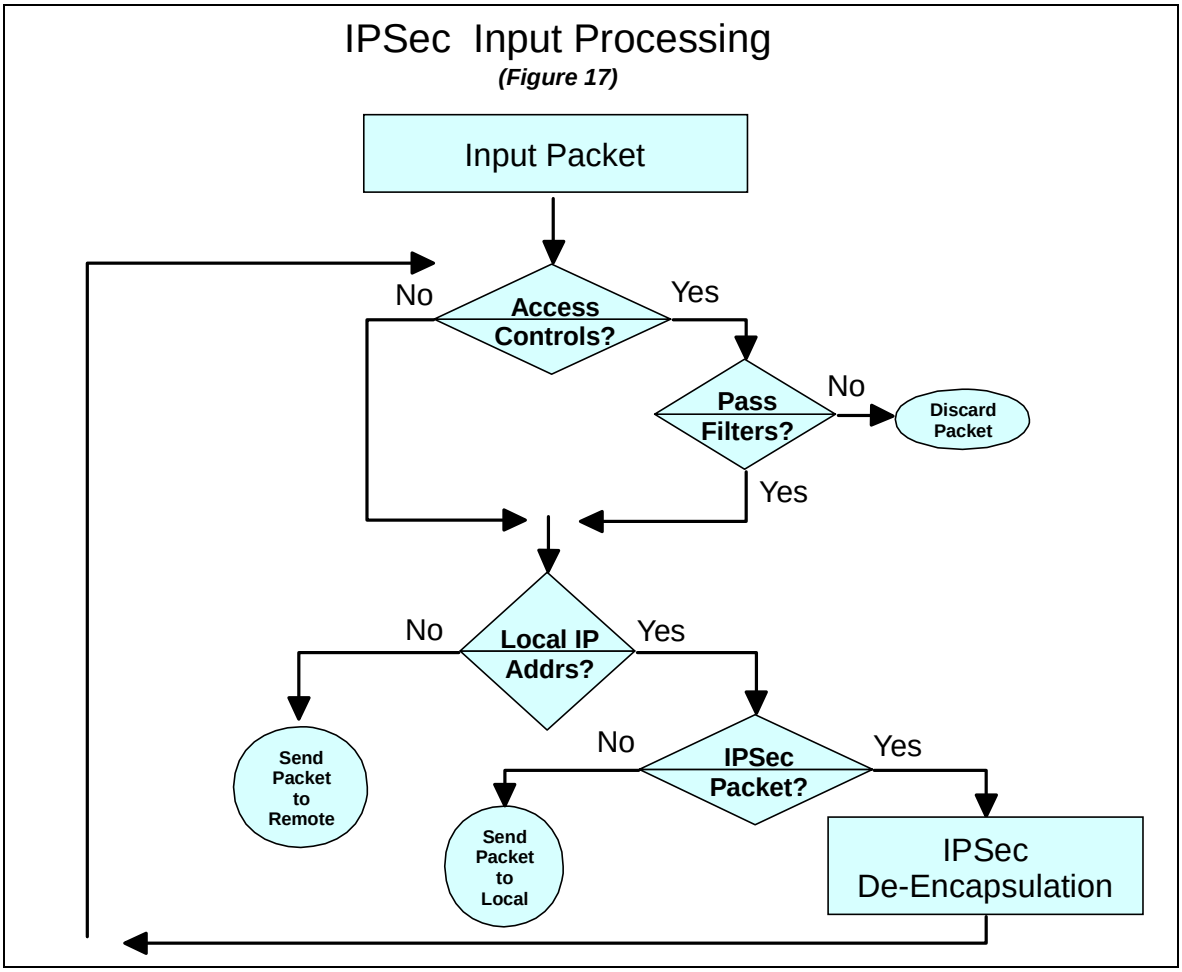


If a port's output filters are enabled, prior to a packet's transmission it is checked against the output filter rules for that port. Packets are identified for placement into a specific IPSec tunnel based on matches on IP addresses and/or protocols as specified by the filter's entries.

Note that packets that will be placed into a tunnel are actually sent though filters twice. The first pass identifies the packets which must be encapsulated and then forwards them to the router's IPSec processing. When the IPSec processing is completed, the encapsulated packets are sent back to the IP stack which routes them (the destination addresses are now tunnel end points) and sends them to the appropriate port and that port's filters. Since they no longer have the original data packet's source and destination IP addresses (but those of the tunnel end points), they fail the "*Put into Tunnel?*" test. A later filter entry to allow the forwarding of traffic with the tunnel end point addresses or IPSec protocol numbers results in the "Yes" path being taken in the generic "*Pass Filters?*" test (actually several tests).

IPSec Input Processing

The following diagram shows the basic flow of port input processing for IPSec.



Only the second test, *"Pass Filters?"*, actually involves the Access Controls filters. All input traffic on a port undergoes one or both of the last two tests. This means that all IPSec processing will take place for local resources and that IPSec packets destined for remote resources will be forwarded **even if Access Controls are not enabled for the interface**. The significance of any security exposures caused by the lack of this extra filtering is debatable.

The only functionality lost without input Access Controls is validation that the de-encapsulated packet should have been placed into the tunnel to begin with. The normal IPSec integrity processing would have detected and dropped any packets which could have somehow been placed into the tunnel by an unauthorized third party. At most, the extra filtering would catch configuration errors made by the router at the other end of the tunnel.

As will be seen next, the benefits of reduced router processing cycles and simpler configuration were judged more important than whatever slight security exposures the inclusion of input Access Controls could address.

IPSec Design

The objectives for the IPSec implementation were:

- **Security**: Privacy and integrity must be provided for user and application data flowing between clients and servers located in branch offices as well as between clients in branch offices and the centrally located mainframes and servers.
- **Efficiency**: Due to the amount of router processing required by IPSec, the access controls had to be as efficient as possible and the amount of data tunneled limited to only what was essential.
- **Ease of Configuration**: Due to the need to use manual tunnels (since that was all that was available at the time) and the large number of routers involved, a simple and standardized Access Control scheme had to be developed.
- **Availability**: The provision of security must be as fault tolerant as the Frame Relay and IP/OSPF infrastructures.

The challenge posed by the last objective, Availability, merits additional discussion.

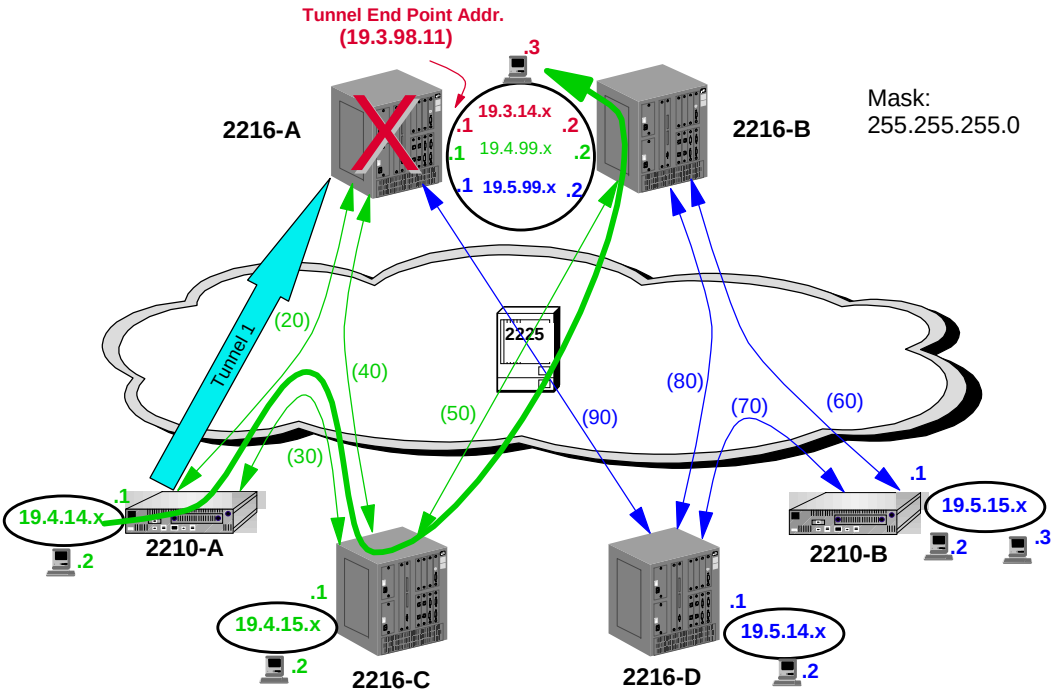
IPSec Availability Challenge

When IPSec tunnels with "Tunnel Mode" are used the actual source and destination addresses of the tunneled IP packet (hereafter referred to as the "data packet") is encapsulated within another IP packet (hereafter referred to as the "IPSec packet") which has the tunnel end points as the source and destination addresses. The tunnel end point addresses, not the addresses of the encapsulated packet, are used to route the packet through to the end point router. As explained earlier the mapping between a data packet's destination address and the correct tunnel end point address is provided ***statically*** via a port's output filters. The challenge with regard to IPSec availability is to continue to provide security when the original tunnel end point is no longer

reachable but the data packet's destination is. The following diagram will help illustrate this scenario.

IPSec Tunnel End Point Failure Scenario

(Figure 18)

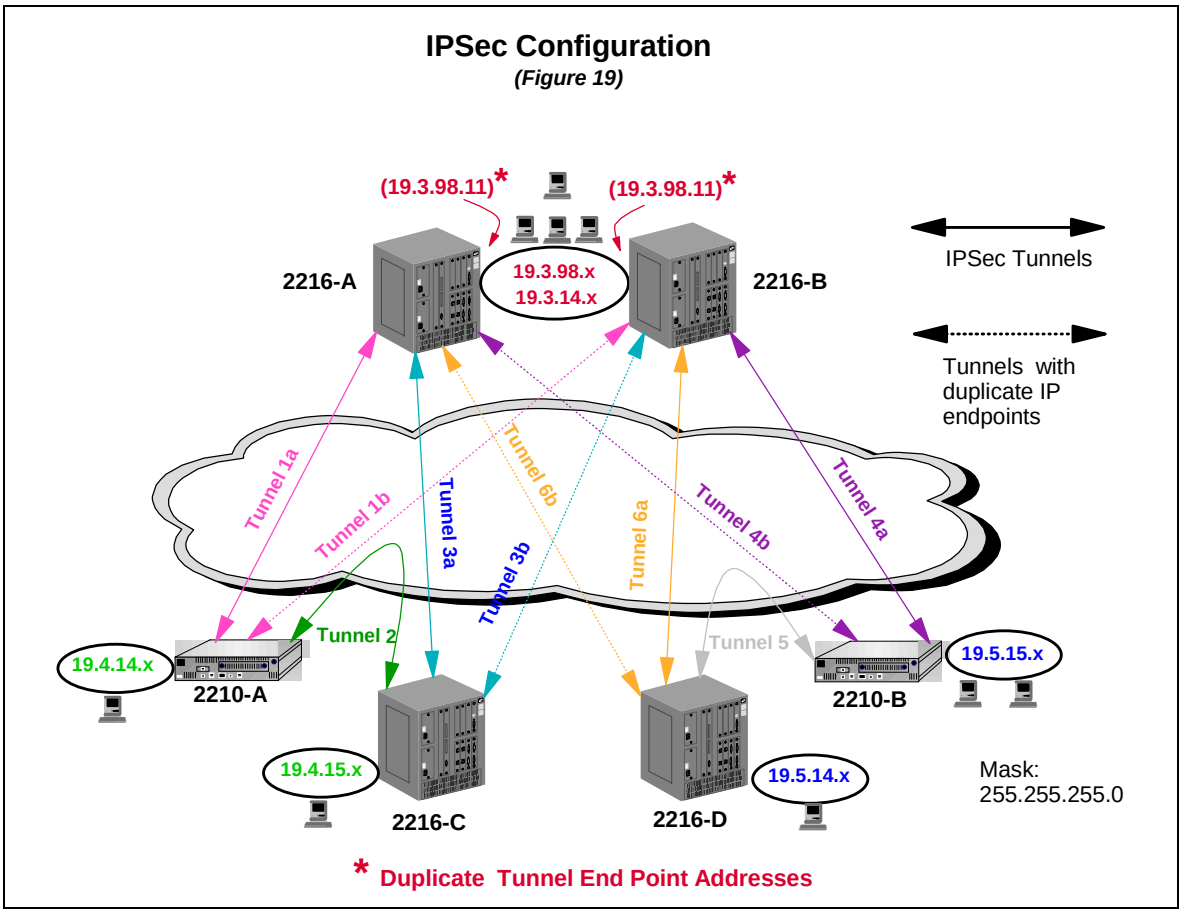


1. The work station attached to 2210-A's token ring is sending packets to a workstation attached to the Central Site token ring which has an IP address of 19.3.14.3. 2210-A normally routes the packets over Frame Relay PVC 20, which resides on its only WAN port. Prior to transmission, the packets are sent to the output filters where they are placed into Tunnel 1 which has a destination tunnel end point in 2216-A of 19.3.98.11 (also reached over PVC 20).
2. 2216-A or its token ring port (on which 19.3.98.11 is defined) become inoperable.
3. 2210-A finds a new path to 19.3.14.3 via 2216-C and 2216-B (as shown by the bold green line)
4. The next packet should go out on PVC 30 but this PVC resides on the same physical interface for which the output filters for the tunnel end point of 19.3.98.11 were configured. Those filters have no way of knowing that Tunnel 1's end point address is no longer reachable so they continue to map all packets directed to 19.3.14.3 into an IPSec tunnel with an unreachable end point address. Per the output filter flowchart (*Figure 16*, page 35) the IPSec packets are then sent to 2210-A's routing processing which knows it can no longer reach the tunnel endpoint address and throws them away ("Destination Unreachable").

The basic problem is that the mapping of a data destination to a tunnel end point is static. There is no mechanism by which the filter can detect when a tunnel end point address is no longer reachable and then pick an alternate end point. The solution to this problem as well as other IPSec implementation details will be covered next.

IPSec Configuration Details

The following diagram shows the IPSec implementation. Some of the Central Site token ring IP addresses have been omitted to simplify the diagram. Tunnel configuration will be discussed first followed by the associated Access Control configurations.



Tunnel Configuration Details

A simple though unorthodox solution to the IPSec tunnel end point redundancy problem was implemented. The same tunnel end point addresses was used in both 2216-A and 2216-B. Though this seems like almost an “unholy” act, no problems will occur as long as that address is only used as a tunnel end point. For example, if 2216-A tried to ping the WAN port on 2216-B using the duplicate address as the source, the ping would get to 2216-B’s port but never return

because 2216-B would consider it a local address. In normal operation, both 2216-A and 2216-B would advertise reachability to 19.3.98.11 and the Branch and Concentration Point routers would pick the one that had the lowest cost path. Since the ultimate "targets" of any traffic were really the stations on the Central Site ring, it really didn't matter which router's tunnel end point address was actually used.

The address was assigned to both routers' token ring ports with OSPF enabled so that the routers would advertise it to the network. The addresses had to be assigned to the token ring ports so that failure of that port would cause the router would stop advertising its reachability forcing the use of the other router's port. **Note that it is very important that *Non-Broadcast* be set to *yes* when configuring OSPF for this address to avoid OSPF related problems caused by duplicate IP addresses on the token ring.**

As shown in *Figure 19* (page 39) there are 3 different types of tunnels in the design:

1. between Branch Office Routers and Central Site Routers (Tunnels 1a/b and Tunnels 6a/b): There are really only two tunnels with duplicate Tunnel End Point definitions on both Central Site Routers.
2. between Branch Office Routers and Concentration Point Routers (Tunnels 2 and 5).
3. between Concentration Point Routers and Central Site Routers (Tunnels 3a/b and Tunnels 4a/b): As with the Branch Office Router tunnels, there are really only two tunnels with duplicate Tunnel End Point definitions on both Central Site Routers.

Note that the tunnels do not have to traverse the Frame Relay PVCs configured directly between the routers. For example, looking at *Figure 18* (page 38), if PVC 30 between 2210-A and 2216-C became inactive the IPSec traffic would be rerouted on PVC 20 through Central Site router 2216-A via the same Tunnel 2 end point addresses.

Traffic between work stations attached to Branch or Concentration Point routers and Central Site hosts/servers will only undergo one encapsulation/de-encapsulation. The same is true for traffic between a Branch Router and its area's Concentration point router. Traffic between Concentration Point routers, between Branch office routers in different areas, and between Branch Office and Concentration Point routers in different areas will undergo two encapsulations and de-encapsulations.

Following are some additional details that may be of interest concerning specific options that were used to configure the routers:

- Tunnel Parameters
  - Lifetime: 0 (minutes) so tunnels never expired
  - Encapsulation Mode: Tunnel
  - Policy: AH-ESP
    - AH Algorithm: HMAC-MD5
    - ESP Algorithm: HMAC-MD5
    - Encryption Algorithm: DES-CBC

- Replay Prevention: No - Replay Prevention is not recommended for manual tunnels since the tunnel's sequence number counter does not wrap (per the IPSec standard). When the counter reached its limit the Security Association (SA) would have to be reestablished.
- Tunnel End Point Addresses
  - Local Tunnel Address:
    - Central Site routers: This is the duplicate tunnel end point IP address which **must** be assigned to the token ring ports on both Central Site routers. Since the servers on the ring are the ultimate targets and sources for the tunneled traffic, a Central Site router must stop advertising reachability to the tunnel end point if its token ring port becomes disabled.
    - Branch office and Concentration Point routers: IP address of the local active LAN port used for normal data traffic
  - Remote Tunnel Address: appropriate IP address of the other end of the tunnel

- Internal Router Address: Though this is not part of the tunnel definitions it is important to the correct operation of the tunnels.
  - Central Site routers: Since the Internal Router Address is advertised as reachable regardless of the state of the interface where the address is assigned, the tunnel end point address **cannot** be used for the Internal Address. It's important that the tunnel end point address not be advertised as reachable if the token ring port goes down so that the other Central Site router's duplicate tunnel end point address can the be used by all the routers in the network.
  - Branch Office and Concentration Point routers: In the test system there was only one LAN port configured for each of these routers. This may not be the case in the customer's network so failure of one LAN port must **not** cause the tunnel end point address from being advertised. Thus, for these routers it's important that the tunnel end point address **is** configured as the Internal Address.
- As mentioned earlier *Non-Broadcast* **must** be set to ***yes*** when configuring OSPF for the duplicate tunnel end point addresses to avoid OSPF related problems on the token ring.

Access Control Configuration Details

Efficiency and Ease of Configuration were other key IPSec related objectives:

- Efficiency: Due to the amount of router processing required by IPSec, the access controls had to be as efficient as possible and the amount of data tunneled limited to only what was essential.
  - Since each of the two Central Site 2216s could be handling up to 200 branch routers (possibly 400 in failure scenarios), limiting the amount of processing done was important. To minimize this overhead, only output controls would be used in all the routers as it was decided lack of input filers would not significantly impact security.
  - Only application traffic to and from user/central-sites token rings would be tunneled. Traffic such as Telnet and SNMP would not be tunneled.

- Ease of Configuration: Due to the need to use manual tunnels and the large number of routers involved, a scheme had to be developed which would minimize the amount of configuration necessary in the Central Site and Concentration Point routers when new branches were added. Similarly, the access controls for the Branch Office routers had to be simple and standardized for easy implementation as new routers were added to the network.

As both the definition and order of each entry is important to a filter's correct operation the next sections provide details on the output filters used in each of the routers in the test system.

Area 0.0.0.1 Branch Office and Concentration Point Router Output Filters

2210-A Output Filters

(Figure 20)

Source IP	Source Mask	Destination IP	Destination Mask	Protocol To/From	Destination Port To/From	Source Port To/From	Access Control	Tunnel ID
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	50/51	0/65535	0/65525	Inclusive	N/A
19.0.17.0	255.0.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A
0.0.0.0	0.0.0.0	19.0.17.0	255.0.255.0	0/255	0/65525	0/65525	Inclusive	N/A
19.4.14.0	255.255.255.0	19.4.0.0	255.255.0.0	0/255	0/65525	0/65525	Inclusive	2
19.4.14.0	255.255.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	1
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A

Following is an explanation of each of the entries in Branch Office router 2210-A's output filters. Examples will be provided later to demonstrate how the filters operate.

- **First Entry**: This allows all ESP (Protocol 50) and AH (Protocol 51) packets (IPSec Packets in which data packets have already been encapsulated) to pass regardless of their source or destination IP addresses. Explicit source and destination tunnel addresses could have been specified but this would have been impractical in a large network since it would have required an entry for every tunnel in the Area (both locally originating tunnels and those for which the router would just be an intermediate transit node). It was important to make this the first entry in the table since the IPSec packets would have (incorrectly) matched later IP address based entries.
- **Second Entry**: One of the ways IPSec processing was minimized was to try to restrict tunneled traffic to only application data traffic. Traffic such as Telnet and SNMP was not to be tunneled if possible. This was accomplished by excluding packets that have either a WAN port source or destination address. The Second entry will detect traffic originating from within a router (ie. a Telnet session to another router) and traffic from this router

sent in response to Telnets or SNMP requests which had been directed to its WAN port's address.
    - A noncontiguous mask was used to identify WAN port IP addresses since, per the addressing scheme used, all had a ".17" in the third octet. This allowed any such traffic from any router in the network to be matched. Similar masking could be done in the actual network implementation which, hopefully, would make more efficient use of the address space than was done in the test system.
    - IP address was used for filtering instead of Port or Protocol to limit the number of entries and so that the entry could be as generic as possible.
- **Third Entry**:
    - This is similar to the Second entry but handles the situation in which the destination, not the source address, is a WAN port. This includes scenarios such as the SNMP GUI configurator (running on a locally LAN attached PC) importing a configuration from a remote router. The source is a LAN address but the destination is a WAN address. Note that the Second and Third entries must precede the Fourth and Fifth entries to keep the traffic from being tunneled.
    - If it becomes necessary to tunnel Telnet or SNMP traffic (ie. configurations) to or from a LAN attached work station all that has to be done is specify a router's LAN port address as the target instead of its WAN port's address. This will result in the "source" for outgoing traffic to be the LAN port's address which will fail this test and match one of the later ones that cause encapsulation. This is only possible from a LAN attached work station because any traffic originating from within the router (ie. an operator logged into 2210-A and telneting into another router) will always use the address of the egress port as the source address. In this network, that will always be a WAN port address which will match the Second entry and not be tunneled.
- **Fourth Entry**: This entry places all packets from this router's LAN's subnet which are destined for any other address in its OSPF area (19.4.0.0) into Tunnel 2 to the area's Concentration Point router (2216-C). It was important to place this entry prior to the next one so that intra-area traffic would, per the design, always use Tunnel 2 instead of the tunnel to the Central Site router.
- **Fifth Entry**: Any data traffic coming from this router's LAN that wasn't already picked off by the previous entry has to be heading outside this OSPF Area so it must use Tunnel 1 to the Central Site. In the IPSec configuration diagram (*Figure 19*, page 39) this would be marked as Tunnel 1a (Tunnel 1b, via 2216-B, would be used if no path to 2216-A was available or if 2216-A's token ring port was inoperable). Note that a wild card of 0.0.0.0 is used for the destination.
- **Sixth Entry**: An all inclusive wild card was used to include all other traffic during testing though hits on the access control entries were reviewed to ensure that the preceding entries were operating as intended. It would be up to the network implementers if they wanted to include this entry as it would result in less control over the tunneling of the traffic.

No changes would have to be made to this table as the result of adding new routers into the network. Also, as you'll soon see, the first three filter table entries are common to **all** routers in the network (Branch Office, Concentration Point, and Central Site). Only the Fourth and Fifth

(highlighted) entries are unique to any particular Branch Office router; the only difference being the LAN IP addresses. This consistency greatly simplifies the implementing of new routers.

To get a better idea of how the entries operate, following are some examples you can use to "walk" through 2210-A's filters. Remember, the fist "match" is always taken.

- Ping from 2210-A's work station to a work station on the Central Site token ring (19.4.14.2 -> 19.3.14.3): This will match the Fifth entry and be encapsulated in Tunnel 1.
- Ping from 2210-A's workstation to a work station on the 2216-C's token ring (19.4.14.2 -> 19.4.15.2): This will match the Fourth entry and be encapsulated in Tunnel 2.
- Ping from 2210-A's work station to a work station on 2210-B's token ring (19.4.14.2 -> 19.5.15.2): This will match the Fifth entry and be encapsulated in Tunnel 1.
- Traffic when a work station on the Central Site token ring Telnets into 2210-A using that router's WAN port's address (19.4.17.3) as the destination: The filter will handle the traffic sent from 2210-A back to the work station (source: 19.4.17.3, destination: 19.3.14.3). Matches will be made with the Second filter entry and the traffic will not be tunneled.
- Traffic when a work station on the Central Site token ring Telnets into 2210-A using the router's LAN port's address (19.4.14.2) as the destination: The outgoing traffic will have a source of 19.4.14.2 and a destination of 19.3.14.3 so matches will occur for the Fifth entry and the packets will be placed into Tunnel 1. Note that this scenario is the same as the previous except that 2210-A's LAN port was used as the Telnet's target instead of its WAN port's address.
- Telnet by an operator with a console connection into 2210-A to 2216-D's LAN port address (source: 19.4.17.3, destination: 19.5.14.1): This will match the Second entry so the traffic will not be tunneled.

Examples won't be provided for the operation of the other router's filters as they all work basically the same.

2216-C Output Filters

(Figure 21)

Source IP	Source Mask	Destination IP	Destination Mask	Protocol To/From	Destination Port To/From	Source Port To/From	Access Control	Tunnel ID
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	50/51	0/65535	0/65525	Inclusive	N/A
19.0.17.0	255.0.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A
0.0.0.0	0.0.0.0	19.0.17.0	255.0.255.0	0/255	0/65525	0/65525	Inclusive	N/A
19.4.0.0	255.255.0.0	19.4.14.0	255.255.255.0	0/255	0/65525	0/65525	Inclusive	2
19.4.0.0	255.255.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	3
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A

Following is the explanation of each entry in Concentration Point router 2216-C's output filter shown above:

- **First Entry**: Same as for 2210-A
- **Second Entry**: Same as for 2210-A
- **Third Entry**: Same as 2210-A
- **Fourth Entry**: This is the reverse of the Third entry in 2210-A. Here, the filter will place the traffic destined **for** 2210-A into Tunnel 2. Again, this entry must precede the next one to force intra-area traffic over the tunnel dedicated for that purpose. An entry like this must be added for each Branch Office router in this Concentration Point Router's OSPF area (or for each different subnet on each Branch Office router if some form of mask summarization couldn't be used).
- **Fifth Entry**: This is similar to 2210-A's entry, since all traffic not destined within Area 0.0.0.1 is sent to the Central Site in Tunnel 3. In the IPSec configuration diagram (*Figure 19*, page 39) this would be marked as Tunnel 3a (Tunnel 3b, via 2216-B, would be used if no path to 2216-A was available or if 2216-A's token ring port was inoperable).
- **Sixth Entry**: Same as for 2210-A

As mentioned before the first three filter table entries are common to **all** routers in the network (Branch Office, Concentration Point, and Central Site). Only the Fourth and Fifth entries are unique to a specific Concentration Point router; the only difference being the IP addresses. The only changes that should have to be made to this table would be the addition of an entry (like the Fourth highlighted entry above) whenever a new Branch Office router was added to **this** OSPF area (A new tunnel would also have to be defined.). These new entries must always be placed prior to the Fifth entry shown above.

Area 0.0.0.2 Branch Office and Concentration Point Router Output Filters

2210-B Output Filters

(Figure 22)

Source IP	Source Mask	Destination IP	Destination Mask	Protocol To/From	Destination Port To/From	Source Port To/From	Access Control	Tunnel ID
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	50/51	0/65535	0/65525	Inclusive	N/A
19.0.17.0	255.0.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A
0.0.0.0	0.0.0.0	19.0.17.0	255.0.255.0	0/255	0/65525	0/65525	Inclusive	N/A
19.5.15.0	255.255.255.0	19.5.0.0	255.255.0.0	0/255	0/65525	0/65525	Inclusive	5
19.5.15.0	255.255.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	4
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A

The only differences between this table and 2210-A's are the IP addresses and tunnel numbers specified in Fourth and Fifth entries.

2216-D Output Filters

(Figure 23)

Source IP	Source Mask	Destination IP	Destination Mask	Protocol To/From	Destination Port To/From	Source Port To/From	Access Control	Tunnel ID
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	50/51	0/65535	0/65525	Inclusive	N/A
19.0.17.0	255.0.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A
0.0.0.0	0.0.0.0	19.0.17.0	255.0.255.0	0/255	0/65525	0/65525	Inclusive	N/A
19.5.0.0	255.255.0.0	19.5.15.0	255.255.255.0	0/255	0/65525	0/65525	Inclusive	5
19.5.0.0	255.255.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	6
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A

The only difference between the above table and 2216-C's are the IP addresses and tunnel numbers specified in Fourth and Fifth entries.

Central Site Router Output Filters

2216-A&B Output Filters

(Figure 24)

Source IP	Source Mask	Destination IP	Destination Mask	Protocol To/From	Destination Port To/From	Source Port To/From	Access Control	Tunnel ID
0.0.0.0	0.0.0.0	19.15.14.0	255.255.255.0	50/51	0/65535	0/65525	Inclusive	N/A
19.0.17.0	255.0.255.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A
0.0.0.0	0.0.0.0	19.0.17.0	255.0.255.0	0/255	0/65525	0/65525	Inclusive	N/A
0.0.0.0	0.0.0.0	19.4.14.0	255.255.255.0	0/255	0/65525	0/65525	Inclusive	1
0.0.0.0	0.0.0.0	19.4.15.0	255.255.255.0	0/255	0/65525	0/65525	Inclusive	3
0.0.0.0	0.0.0.0	19.5.15.0	255.255.255.0	0/255	0/65525	0/65525	Inclusive	4
0.0.0.0	0.0.0.0	19.5.14.0	255.255.255.0	0/255	0/65525	0/65525	Inclusive	6
0.0.0.0	0.0.0.0	0.0.0.0	0.0.0.0	0/255	0/65525	0/65525	Inclusive	N/A

As they must each be capable of providing IPSec services for all Branch and Concentration Point routers in the network (in case the other fails), the output filter tables for both Central Site routers must be identical.

- **First Entry**: Same as all other routers.
- **Second Entry**: Same as all other routers.
- **Third Entry**: Same as all other routers.
- **Fourth Entry**: Places all traffic destined for Branch Office router 2210-A into Tunnel 1.
- **Fifth Entry**: Places all traffic destined for Concentration Point router 2216-C into Tunnel 3. Entries for all Branch Office routers for this Concentration Point router's area must be placed before this entry.
- **Sixth Entry:** Places traffic destined for Branch Office router 2210-B into Tunnel 4.
- **Seventy Entry:** Places all traffic destined for Concentration Point router 2216-C into Tunnel 6. Entries for all Branch Office routers for this area must be placed before this entry.
- **Eighth Entry**: Same as last entry in the tables in all routers.

New entries would be added to the tables when either a new Concentration Point or Branch Office router was added to the network. This also implies new IPSec tunnels would be defined for them. These new entries would be grouped by Area just as the fourth and fifth entries are for Area 0.0.0.1 and the sixth and seventh entries are for Area 0.0.0.2. As noted above, the last entry for each Area's group of entries must pass the traffic to the Concentration Point router for that Area. Actually, all that's important is that all the Branch Office router entries for an area precede the entry for the Concentration Point router. The entry for the Concentration Point router does not have to immediately follow the entry of the last Branch Office router of the area though this might be a good idea from the perspective of "dividing" the filters into distinct sections for organization/documentation purposes.

IPSec Configuration Testing Results

No problems were experienced with the configuration or operation of the IPSec tunnels. The use of duplicate tunnel end point addresses on the Central Site routers worked perfectly in all steady state and failure scenarios. On the other hand, due to the performance and ease of configuration requirements getting the output filters correct required much trial and error. Only the "final versions" were discussed.

Though things worked well, a few additional IPSec related comments are in order:

- At the time of this testing only manual IPSec Tunnels were available. Due to the large number of routers involved in the actual network and the effort/exposures associated with managing such a large number of keys, automatic key management provided by ISAKMP/Oakley based functionality should be considered when it is available. Further testing would be needed to validate that a system using automatic key processing would be able to provide the same level of redundancy as does the manual tunnel configuration. For continued use of manual tunnels, a good procedure for regular key changes is crucial for ensuring the security of the system.
- Though DLSw was not included in the testing it is recommended that IPSec Tunnel Mode be used for all traffic. Though use of Transport Mode for DLSw traffic would save some WAN bandwidth (due to smaller IPSec header since source/destination addresses are not encapsulated) it could almost double the number of tunnels which had to be supported by the Central Site routers greatly increased the load on the CPU. It would so add an additional degree in configuration complexity (more keys, additional filter table entries, etc.) to all the routers.
- Though OSPF load balances traffic when equal cost paths to destinations exist, there are two different methods that can be used:
    - *Per Destination Multipath*: When there are multiple equal cost paths to destinations, the router will balance the traffic **per destination**. For example, assume *Path 1* and *Path 2* have the same cost and both could be used to reach *Destination A* and *Destination B* in one hop. The router would use *Path 1* for **all** traffic to *Destination A* and *Path 2* for **all** traffic to *Destination B* (or visa versa). *Per Destination Multipath* is the default configuration value and was used in this test.
    - *Per Path Multipath*: When enabled and there are multiple equal cost paths to **a** destination, the router chooses the path for forwarding each packet in a round-robin fashion. This balances the network utilization of paths to **a** specific destination by using a different path to transmit **each** packet.

  Now remember in this network, all traffic destined for a station attached to the Central Site Token Ring is encapsulated in an IPSec packet with the **same** tunnel end point destination address, 10.3.98.11 (see *Figure 19*, page 39). This means that no load balancing was really being done for traffic passing through either of the Concentration Point Routers (2216-C and 2216-D) since *Per Destination Multipath* only load balances for **different** destinations (see Figure 7, page 23). A detailed traffic flow analysis will have to be done to determine if load balancing is really appropriate for this network

given some of the other requirements. Following are recommendations pending the results of that study:

- Load Balancing: *Per Path Multipath* must be configured on each Concentration Point Router.
- No Load Balancing: If the Concentration Point Routers' dual PVCs to the Central Site are to be considered *primary* and *secondary* then use the default *Per Destination Multipath* **and** set the OSPF Cost for the *primary* PVCs lower than that of the *secondary* PVCs in the Concentration Point Routers. In the test system this means setting the cost of PVCs 50 and 90 (*secondary*) higher than that of PVCs 40 and 80 (*primary*) (see Figure 7, page 23). This is necessary so that the *primary* PVC is used even if the Concentration Point Router learns about the *secondary* first or after the failure and recovery of one of the Central Site Routers. For example, if all the above PVCs had equal costs and Central Site Router 2216-A failed, 2216-C would use PVC 50 for all traffic to the Central Site. Unfortunately, when 2216-A came back on line, 2216-C would continue to use PVC 50 since it had the same cost as *primary* PVC 40. Note that these recommendations are only applicable to traffic flowing **from** the Concentration Point Routers **to** the Central Site. The routes taken by traffic flowing in the reverse direction in the actual network will be dictated by which router contains the Default Gateway being used by the responding server. In this test network all the workstations attached to the Central Site Ring used 2216-A as their Default Gateway but in the actual network DFG responsibilities will probably be divided between the two Central Site routers.

- Though no capacity/performance problems were experienced or anticipated, no performance analysis was done during this testing. Possible impacts on router CPU utilization and response times (especially in the scenario when one of the central site routers is lost) requires further study.

Summary

As stated in the *Preface*, the overall objective of the testing described in this document was to validate that the proposed design would meet the customer's objectives for connectivity, availability, security, efficiency, and ease of configuration. As the result of the testing a problem with the planned OSPF configuration was uncovered and a practical alternative validated. Other elements of the proposal were validated and a number of configuration details were identified. Though this should make the network's implementation much simpler, additional testing to address other design elements and concerns (DLSw, dial backup, performance, etc.) is needed before the roll out can take place.

The objective of this document was to share the testing experiences with the reader so that time and effort can be saved by using some of the recommended techniques and also by avoiding some of the problems encountered during the test. I can only hope it will meet that objective.

Appendix A: Test Procedures

Following are the procedures used to validate the complete configuration (Frame Relay, IP/OSPF, VRRP, and IPSec). Tests were run in the order shown; check marks being placed on the provided lines when successful or notes added to document problems. It was important to rerun all the tests whenever changes were made to the OSPF configuration or Access Control filters as any modification (no matter how seemingly minor) could have major impacts on operation.

1. Verify non end station date not put into tunnels (clear all tunnels first)
   a. Telnet from Routers (verify no tunnel traffic):

      ___
      i. 2216-A -> 2210-A (19.4.14.1): ___
      ii. 2210-A -> 2216-C (19.4.15.1): ___
      iii. 2216-C -> 2216-C (19.5.14.1): ___
      iv. 2216-C -> 2210-B (19.5.15.1): ___
      v. 2210-B -> 2216-B (19.3.14.2): ___
      vi. Logout back to 2210-A: ___
      vii. 2210-A -> 2216-C (19.5.14.1): ___
      viii. Logout back to 2210-A: ___
      ix. 2210-A -> 2216-B (19.3.14.2): ___
   b. Telnet from WS (Work Station) on 19.3.14.x (verify no tunnel traffic):

      ___
      i. 2210-A (19.4.17.3): ___
      ii. 2216-C (19.4.17.4): ___
      iii. 2216-C (19.5.17.3): ___
      iv. 2210-B (19.5.17.4): ___
      v. 2216-B (19.4.17.2): ___
      vi. 2216-A (19.4.17.1): ___
   c. Retrieve configs into WS on 19.3.14.x (verify no tunnel traffic):

      ___
      i. 2210-A (19.4.17.3): ___
      ii. 2216-C (19.4.17.4): ___
      iii. 2216-B (19.4.17.2): ___
      iv. 2216-A (19.4.17.1): ___
      v. 2216-C (19.5.17.3): ___
      vi. 2210-B (19.5.17.4): ___
2. Test Tunnels (verify counters incrementing, reset tunnels after each test, all traffic consists of 3 PINGs): ___

a. Tunnel 1: 2210-A WS -> 2216-A WSa (19.3.14.3)
   i. 2216-A: ___
   ii. 2210-A: ___
b. Tunnel 2: 2210-A WS -> 2216-C WS (19.4.15.2)
   i. 2210-A: ___
   ii. 2216-C: ___
c. Tunnel 3: 2216-C WS -> 2216-A WSa (19.3.14.3)
   i. 2216-C: ___
   ii. 2216-A: ___
d. Tunnel 4: 2210-B WS -> 2216-A WSa (19.3.14.3)
   i. 2210-B: ___
   ii. 2216-B: ___
e. Tunnel 5: 2210-B WS -> 2216-C WS (19.5.14.2)
   i. 2210-B: ___
   ii. 2216-C: ___
f. Tunnel 6: 2216-C WS -> 2216-A WSa (19.3.14.3)
   i. 2216-C: ___
   ii. 2216-A: ___
g. Multiple Tunnel Traffic:
   i. 2216-C WS -> 2210-A WS (19.4.14.2)
      1) 2216-C Tunnel 6: ___
      2) 2216-A:
         a) Tunnel 1: ___
         b) Tunnel 6: ___
      3) 2210-A Tunnel 1: ___
   ii. 2210-A WS -> 2210-B WSa (19.5.15.2)
      1) 2210-A Tunnel 1: ___
      2) 2216-A Tunnel 1: ___
      3) 2216-B Tunnel 4: ___
      4) 2210-B Tunnel 4: ___
   iii. 2216-C WS -> 2216-C WS (19.4.15.2)
      1) 2216-C Tunnel 6: ___
      2) 2216-A
         a) Tunnel 3: ___
         b) Tunnel 6: ___

3. Failure Scenarios
   a. Start all following traffic (continuous pings):
      i) 2216-A WSa -> 2210-A WS (19.4.14.2): ___ ___
      ii) 2216-A WSb -> 2216-C WS (19.4.15.2): ___ ___

iii) 2216-A WSc -> 2216-C WS (19.5.14.2): ___ ___
iv) 2216-A WSd -> 2210-B WS (19.5.15.2): ___ ___
v) 2210-A WS -> 2216-C WS (19.4.15.2): ___ ___
vi) 2216-C WS -> 2216-C WS (19.5.14.2): ___ ___
vii) 2216-C WS -> 2210-B WS (19.5.15.2): ___ ___
viii) 2210-B WSa -> 2210-A WS (19.4.14.2): ___ ___
ix) 2210-B WSb -> 2216-C WS (19.4.15.2): ___ ___

b. Deactivate/Reactivate DLCIs: note any traffic termination/verify tunneling
   i. DLCI 20: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   ii. DLCI 30: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   iii. DLCI 40: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   iv. DLCI 50: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   v. DLCI 60: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   vi. DLCI 70: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   vii. DLCI 80: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
   viii. DLCI 90: ___
      1) Deactivate DLCI: _____
      2) Reactivate DLCI: _____
c. 2216-A V.35 Failure
   i. Pull Cable from Router: ___
      ii. Telnet from 2210-A into all other routers: ___
      iii. Telnet from 2210-B into all other routers: ___
   iv. Replace Cable: ___
d. 2216-B V.35 Failure
   i. Pull Cable from Router: ___

ii. Telnet from 2210-A into all other routers: ___
iii. Telnet from 2210-B into all other routers: ___
iv. Replace Cable: ___

e. 2216-A TR Port Failure
i. Pull Cable from Router: ___
ii. Telnet from 2210-A into all other routers: ___
iii. Telnet from 2210-B into all other routers: ___
iv. Replace Cable: ___

f. 2216-B TR Port Failure
i. Pull Cable from Router: ___
ii. Telnet from 2210-A into all other routers: ___
iii. Telnet from 2210-B into all other routers: ___
iv. Replace Cable: ___

g. 2216-A V.35/TR Ports (Simulates 2216-A going down):
i. Pull Cable from Router: ___
1) Telnet from 2210-A into all other routers: ___
2) Telnet from 2210-B into all other routers: ___
3) Retrieve configs from all other routers: ___
ii. Replace Cable: ___

h. 2216-B V.35/TR Ports (Simulates 2216-B going down):
i. Pull Cable from Router: ___
1) Telnet from 2210-A into all other routers: ___
2) Telnet from 2210-B into all other routers: ___
3) Retrieve configs from all other routers: ___
ii. Replace Cable: ___

Appendix B: Recommended Reading

- Kearby, T., Boelaars, S., Lingafelt, C., and Samra, K., *A Comprehensive Guide to Virtual Private Networks, Volume II: IBM Nways Router Solutions*, IBM, Research Triangle Park, NC, 1998
- Moy, John T., *OSPF: Anatomy of an Internet Routing Protocol*, Addison-Wesley Longman, Inc., Reading, MA, 1998
- Smith, Richard E., *Internet Cryptography*, Addison-Wesley Longman, Inc., Reading, MA, 1997

