FIA DRP Phase I: Network Infrastructure

Project Report for the Capstone Project

Leave as Blank

Contents

1. Overview

This report gives a detailed description of the initiation, design and implementation of the development of a Disaster Recovery Plan (DRP) for a major Financial Service Institution. For confidentiality reasons, and for brevity, we refer to this institution as Financial Institution Alfa (FIA) in this report. Specific references to FIA (e.g. additional context) are available on request.

1.1 Business Requirement

Disasters are hard to predict and all have different causes and reasons. To lessen the impact of disasters, a well thought Disaster Recovery Plan (DRP) of its IT systems is becoming mandatory for modern enterprises.

FIA is a provincial crown company in Canada. It provides a range of financial services for farmers which includes insurance, loan and farm income disaster assistance.

Business Continues Plan (BCP) and Disaster Recovery Plan (DRP) play critical roles for FIA's success and survival. There are several reasons to have a good DRP plan for FIA:

- As a Crown Corporation
  FIA needs to meet the provincial DRP requirements as a crown corporation.
- As a Financial Organization
  FIA is obliged to follow a set of regulations.
- As a profit-driven enterprise

  Without a DRP, it will directly result in revenue loss when disaster happens.

However, FIA is currently using the DRP which was first written in 2005. No review or improvement has been carried out since then. Furthermore, there are several new business systems which have been built without a DRP such as the finance system etc.

As of that, FIA needs to review and improve a valid DRP for critical business systems and operation. In some cases, it is required to redesign the DRP.

This document focuses on the DRP for FIA's network infrastructure which is the fundamental supporting platform for the organization's business systems.

1.2 Project Scope

The FIA DRP is formulated to be designed over several, but incremental phases. The first phase of this project is referred to as '**Phase I'** by the project team, and is defined in concert with FIA's business units. This document presents the major objectives and Phase I deliverables.

DRP Site

There is a separate DRP site located at City B. This DRP site is about 100 kilometers away from the primary site in City A. In accordance to the local building restrictions and the associated environment at City B, this DRP site does not have the same level of

resilience as the primary site. Factors affecting "resilience" include environmental factors such as the physical datacenter room size, physical security access controls, power, and the availability of supporting technicians. This site is used as cold site in the DRP Phase I which means that there is no critical business traffic through the systems in this site during normal operation periods.

Disaster Declaration Condition

Disaster declaration condition includes a set of conditions under which the DRP process is to be activated. In FIA DRP phase I, the DRP process needs to be activated in either of the two situations described below:

- The primary site in City A is not able to provide any service to its online customers or its remote office employees in a specific time range (the time range is not clearly defined by FIA). At a point where the FIA business team measures the impact, deems it significant, it will request that the DRP process is activated. For example, an Internet access outage in City A occurred and has lasted for longer than one hour without clear expected deadline to recover.
- The primary site in City A is totally destroyed by natural or man-made disasters. For example, the primary site in City A is wiped out by fire or hurricane.

Recovery Point Object (RPO) and Recovery Time Object (RTO)

FIA has a number of business services that are supported by different IT systems. There is a challenge on resources and timeline to have a well-defined DRP solution for all critical businesses. Due to the complexity that will be required to align all the competing business services to a functional recovery plan, it is necessary to further divide recovery timelines for each of the services between phases.

In FIA DRP Phase I, the targets for RPO and RTO are as follows:

- FIA Finance System achieves 2 days of RTO and 1 week of RPO
- FIA Exchange Mail system achieves 4 hours of RTO and 24 hours of RPO
- FIA WEB Server achieves 4 hours of RTO and undefined RPO

2. Network Requirements Analysis

2.1 DRP System Components and Relation

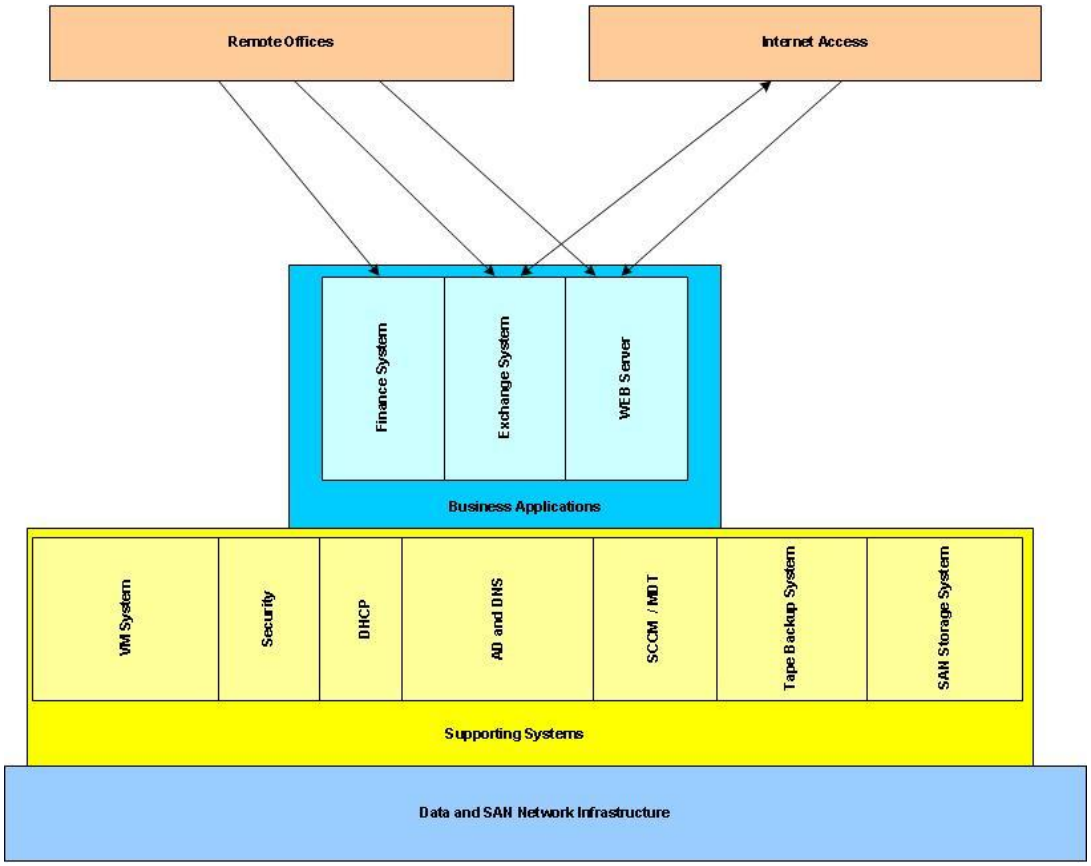

Figure 2-1: DRP System Components and Relation

FIA DRP Phase I includes three critical business systems and seven supporting systems. The three business systems include Finance system, Exchange system and Web server. The seven supporting systems are VM (Virtual Machine) system, Security, DHCP (Dynamic Host Control Protocol), AD/DNS (Active Directory/Domain Name System), SCCM/MDT (System Center Configuration Manager / Microsoft Deployment Tool), Tape Backup system and SAN (Storage Area Network) Storage system.

The Network Infrastructure is required to provide capacity and the requisite density for these systems. It is also required to provide branch office access and connectivity. In section 2.2, we look at each system separately and analysis the network requirements of each system.

2.2 Business Application System Requirements

2.2.1 Finance System

FIA's finance system is based on the commercial applications developed by software company SAP. Currently, all servers of the finance system reside in the primary datacenter located in City A.

In FIA DRP phase 1, FIA business requires that finance system to construct a DRP solution to meet 2 days of RTO and 1 week of RPO. The finance system team is required to create a DRP plan for the finance system. In this plan, a separate finance system infrastructure is to be built on the DRP site (City B). There is no redundant cluster structure in databases and applications. It is a simplified version of the system in City A. The infrastructure composes of:

- 1 physical Database Host
- 1 physical Central Instance Host (Application)
- 2 Virtual Machines (VMs) which host Dialog Instances

DRP Process

Prior to the disaster, the finance DRP system is disabled. When a disaster occurs, the DRP database is to be restored from the weekly backup tapes, and then FIA workstations will be pointed to the new DRP system.

Hardware Requirements

The physical servers from Hewlett-Packard (HP) Company are the dominant servers in FIA datacenter. The FIA DRP finance system needs the following physical devices:

- 1 blade server for database
- 1 blade server for Central Instance
- 1 blade server for one VMware host which hosts two VMs

All of the physical servers of DRP finance system are blades which reside in one HP BladeSystem c7000 enclosure.

Other than servers, there are two specifically purposed printers which are required in the DRP system. These printers will be located in the office area in the disaster site. The printers perform the following functions:

- Check printing
- Batch job printing

Data Network Connection Requirements

As all physical servers of the DRP finance system reside in one enclosure (HP BladeSystem c7000), this enclosure provides the network connection for all blade servers via the enclosures own switching subsystems. For DRP's finance system, each blade needs two separate data network connections to two enclosure switches in order to prevent link failure. The type of these connections is Gigabit Copper Ethernet.

Both printers require Gigabit Copper Ethernet connection.

Storage Network Connection Requirements

Similar to the data network connection, each blade for the DRP finance system requires two separate storage network connections to two enclosure switches to prevent link failure. The storage network provides network storage services for DRP finance system. The type of those connections is Gigabit Copper Ethernet.

Internet Access Requirements

There is no Internet access required for the DRP finance systems. All traffic is internal and connectivity established utilizing the FIA intranet.

Remote Office Access Requirements

Under normal operation status, there is no access from any remote offices excluding management traffic from City A.

During the disaster period, all FIA remote offices require connectivity via specific client software, to the DRP Central Instance.

2.2.2 Exchange System

FIA's electronic mail system is built on Microsoft Exchange 2010 environment. In City A's datacenter, there are three physical servers with an attached local storage system located in City A's datacenter. Each server hosts three Exchange roles: CAS (Client Access Server), Hub Transport and Mailbox Database. These three servers are clustered together to provide a robust email system:

- For client access sessions: two hardware-based load balancers are deployed to distribute client sessions among the three servers.
- For mailbox database high availability: These three servers form a DAG (database availability group). When one server of the DAG has any changes on the mail databases, it automatically replicates the changes to other two servers within the same DAG.

To provide the site resiliency for mailbox database, Microsoft Exchange 2010 recommends extending the DAG to multiple sites. In FIA DRP Phase I, FIA Exchange Team extends the current DAG to City B datacenter. There will be one new server deployed in City B datacenter. This server hosts the same three exchange roles: Client Access Server, Hub Transport and Mailbox Database. This server is clustered with the three servers in City A to provide Exchange DRP system:

- For client access sessions: City B has only one server which has Exchange Client Access Server. This server will not accept any client during the normal operation period. When disaster happens in City A, all the client access sessions are to be redirect to the server in City B. This change is achieved by manually updating the DNS entries of the mail server.
- For mailbox database high availability: The server in City B datacenter is added to the same DAG as the three servers in City A datacenter. This server will

synchronize the mailbox databases with other three servers automatically. The changes on the three servers will be continuously replicated to the server in disaster site.

DRP Process

When primary datacenter fails, three servers in City A will be unavailable and lead to the failure of the DAG (cluster). The following diagram describes the primary datacenter failure scenario.

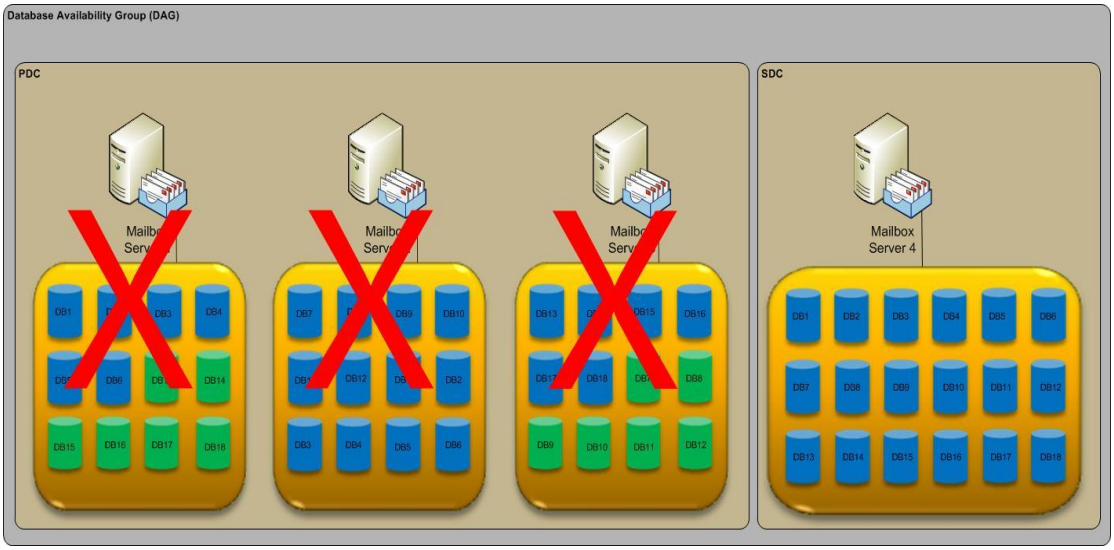

Figure 2-2: Exchange 2010 Fail Over

Note: PDC stands for Primary Data Center which is in City A; SDC stands for Secondary Data Center which is in City B

Hardware Requirements

The Exchange DRP system needs the following devices:

- 1 Exchange Server with directly attached storage
- 1 Microsoft Thread Management Gateway (TMG) in DMZ (Demilitarized Zone) for connection from the Internet
- 1 Microsoft Edge server in DMZ for processing incoming and outgoing SMTP services

Data Connection Requirements

The Exchange server needs two separate data network connections: one for regular data network connection, and the other for replication network. The members of one DAG need to replicate data across each other.

There are specific network performance requirements for replication network which are provided by FIA Exchange team as follows:

- Bandwidth Requirement: The average bandwidth should be over 8Mbps (Megabits per second) and the peak bandwidth should be up to 16Mbps between any two members of the DAG.

- Round-Trip Latency Requirement: The network round-trip time should be less than 400 milliseconds between any two members of the DAG.

SAN Connection Requirements

As the Exchange server has directly attached storage system, there is no need for a connection to the SAN.

Internet Access Requirements

During normal operation period, there are is Internet traffic to or from City B Exchange server. After disaster happens, City B Exchange server has two types of Internet traffic: SMTP (Simple Mail Transfer Protocol) and HTTPS (Hypertext Transfer Protocol Secure). SMTP traffic is to deliver and receive email from outside SMTP servers. HTTPS is for other Exchange services such as Outlook Web Access, Outlook Anywhere, and Outlook ActiveSync.

Remote Office Access Requirements

All FIA branch offices need to access City B's Exchange DRP server after disaster.

Microsoft Outlook clients in remote offices will be redirected to City B Exchange DRP server to retrieve and send email.

2.2.3 Web Server

FIA's website (www.fia.ca) is a critical service to external clients. In FIA DRP phase 1, only static web pages will be provided to clients during the recovery period. Currently, FIA has one backup web server located in City B. This server will be used as the DRP web server in phase 1. This server hosts the static web pages in local disk.

DRP Process

During normal operation time, web administrator manually synchronizes this DRP server with the primary server including:

- Windows OS patches
- Web service configuration
- Static web pages

During disaster period, the recovery process should follow these steps:

- External Firewall administrator confirms that the public IP used by the backup site in City B is redirected to the DRP web server
- External DNS administrator updates the DNS entry to point the www.fia.ca domain to the public IP in City B

Hardware Requirements

The existing standalone server will be used for phase 1.

Data Connection Requirements

This server requires one 100 Mbps port to City B's DMZ switch (this is already provisioned).

SAN Connection Requirements

There is no requirement for SAN connectivity.

Internet Access Requirements

There are two types of Internet access requirements:

- Internet access for Windows OS patches
- After disaster, this DRP web server needs to be accessed by external clients from Internet using http and https.

2.3 Supporting System Requirements

2.3.1 VM System

FIA's VMs (virtual machines) run on VMware's vSphere platform. In DRP Phase I, VMware vSphere is also used to provide VMs in disaster site. The VM system in City B is used to support the multiple guests for Finance system, File/Printer server, SCCM/MDT server and virtual desktops.

In City B, there is one VMware vCenter Server to manage all VMs and hosts. In DRP phase 1, vCenter server in City B has the following responsibilities:

- vCenter server in blade server manages 2 VM hosts
- The VM hosts are alive at all time.
- vCenter server in City B should not synchronize with vCenter server in City A
- Rebuild process will be used if any of the VM Hosts or Guests crash.

Hardware Requirements

Two blade servers are needed to setup two VM hosts.

The following VMs will be created in two hosts:

- 2 VMs for dialog instance of Finance system
- 1 VM for vCenter V 5.0 application
- 1 VM for vCenter database
- 1 VM for File / Print server
- 1 VM for SCCM / MDT
- 2 VMs for virtual desktops

Data Connection Requirements

Each of the VM Hosts requires two 1 Gbps data network connections. For flexibility and scalability, these two connections should be configured as Dot1Q trunk link. There are three different types of IP address requirements for each VM Host:

- VM host management IP
- IP of guest servers on VM host
- IP of virtual desktops on VM host

SAN Connection Requirements

Each of the VM Hosts requires two 1 Gbps SAN network connections. One link should connect to VLAN 10 in SAN network. The other link should connect to VLAN 20 in SAN network.

Internet Access Requirement

VMware vCenter application requires internet access for periodic patch updates before and after disaster.

2.3.2 Security System

In FIA DRP Phase 1, there are no additional security requirements in City B. The current external firewall and DMZ firewall will be the main security guard to protect City B's local network against the Internet. During normal operation, there is no Internet traffic to and from City B's Internet link. After disaster occurs, DMZ firewall and external firewall need to provide security filtering control for the Internet traffic through City B's Internet link.

2.3.3 DHCP System

DHCP (Dynamic Host Control Protocol) is a basic service for end devices. With DHCP service in place, the end devices can automatically configure their network settings according to the information obtained from the DHCP server.

FIA already has one standalone Windows 2003 server as DRP DHCP server. This DHCP DRP server provides DHCP services for City B's local offices and all FIA branch offices. This DHCP is alive at any given time. There is no data synchronization between DHCP servers in City A and City B.

In FIA DRP phase I, no change is necessary to the existing DHCP system.

2.3.4 AD and DNS

Microsoft Active Directory (AD) service provides a central repository and location to manage and authenticate Windows computers and user accounts in FIA. Currently, two domain controllers are located in City B. These two servers synchronize with the two domain controllers in City A. The four domain controllers form one cluster to provide Active Directory service for FIA production environment. FIA internal DNS is integrated with Active Directory. These four domain controllers also provide DNS services.

In FIA DRP phase I, there are no changes to the existing AD and DNS system.

2.3.5 SCCM and MDT

SCCM (System Center Configuration Manager) and MDT (Microsoft Deployment Tool) are utilized after disaster. SCCM provides the patches and updates automatic installation for servers in City B after disaster. MDT provides image deployments for desktop machines in City B after disaster.

Disaster Process

In the pre-disaster stage, SCCM and MDT in City A take the responsibility to install patches and deploy desktop images. The patches files and various image files of SCCM and MDT tool are backed-up to the tape once a week. These tapes are stored in a regional office in City C.

When disaster happens, these tapes are shipped to City B and restored to the SCCM and MDT system in City B.

Hardware Requirements

SCCM and MDT are deployed within one virtual server in City B.

Data Connection Requirements

Since both SCCM and MDT run on a virtual server, data connection is achieved through the virtual connection provided by the host server.

SAN Connection Requirements

The host server of SCCM and MDT provides virtual SAN connection.

Internet Access Requirement

SCCM needs Internet access to obtain the latest Windows OS patches through Microsoft's website after disaster.

Remote Office Access Requirement

MDT needs to be able to build desktop machines in City B's local offices. Administrators of SCCM and MDT require management access in both City A and City B.

2.3.6 Tape Backup System

Backup Server and Tape Library play an important role during DRP recovery process. It provides data backup and restore services for all DRP systems in City B. In DRP phase 1, the main system relies on backup server and tape library is Finance System.

Finance System: The entire production database needs to be restored from tape after disaster. In general, production databases are backed up to tape from production system in City A. These tapes will then be shipped and stored in City C (a regional Office). When disaster occurs, the latest tape will be shipped to City B DRP site. The backup server and tape library provides the data to restore the production databases of appropriate DRP system.

Hardware Requirements

There is one standalone Windows server working as a backup server. Tape Library is one standalone device.

Data Connection Requirements

The backup server needs one 1 Gbps data network connection. Tape Library connects to backup server through SCSI (Small Computer System Interface) interface. Tape Library doesn't need data network connection.

2.3.7 SAN Storage System

In FIA DRP phase 1, one NetApp SAN3140 with 24 disk shelves provides the storage service for the following functionalities before and after disaster happens:

- Virtual servers and virtual desktops
- Additional storage space for file server
- Storage space for Finance system database

There is no data replication requirement between NetApp SAN3140 in City B and SAN filers in City A's datacenter. NetApp SAN 3140 only provides iSCSI (Internet Small Computer System Interface) connections for local servers.

Hardware Requirements

One standalone appliance NetApp SAN 3140 works as storage filer to provide storage services.

Data Connection Requirements

NetApp SAN3140 requires the following network connections:

- One 100 Mbps connection for management access
- One 100 Mbps connection for support technician to perform troubleshoot

SAN Connection Requirements

NetApp SAN3140 requires the following SAN connections

- One 1 Gbps connection to VLAN 10 in City B's SAN network
- One 1 Gbps connection to VLAN 20 in City B's SAN network

Internet Access Requirements

There are two types of Internet access requirements before and after disaster happens:

- NetApp SAN3140 needs to send the monitoring log to vendor through management port
- When NetApp SAN3140 malfunctions, vendor needs to access the SAN3140 through the troubleshooting port

3. Design

3.1 Network Topology Design

3.1.1 Current City B Network Topology

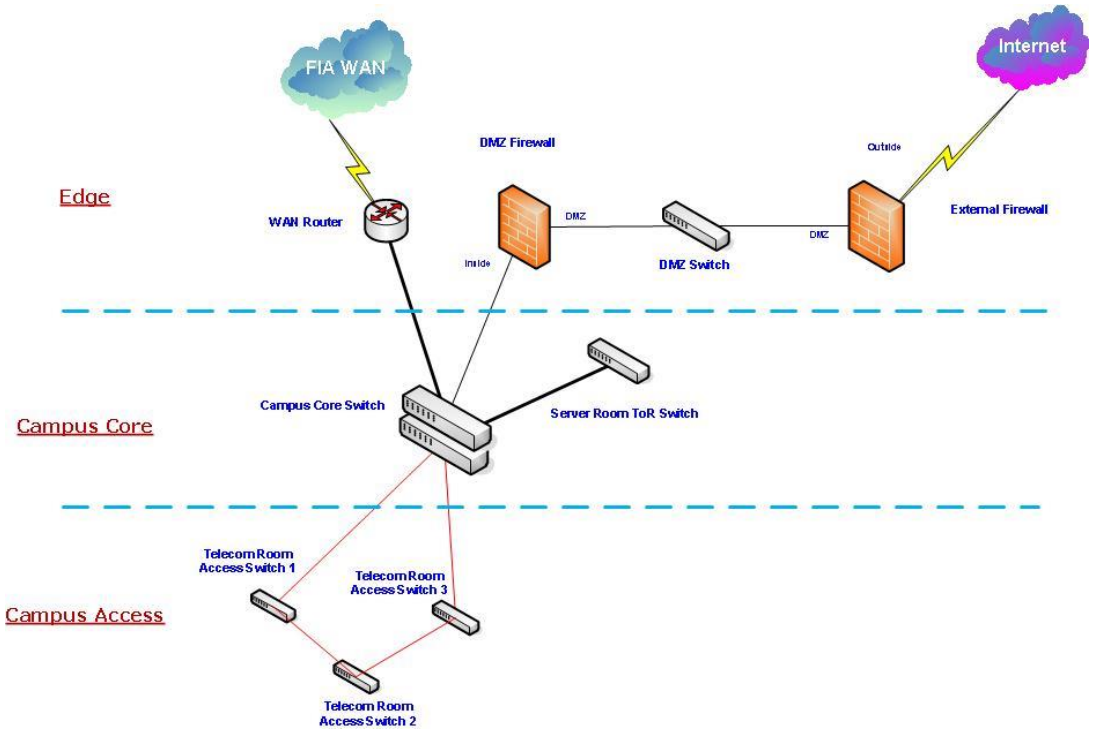

Figure 3-1: City B's Current Network Logical Topology

The current City B DRP site network can be divided into three parts: Edge network, Campus Core network and Campus Access network.

In the Edge network, the WAN (Wide Area Network) Router is connected to Service Provider CPE (Customer Premises Equipment). It provides the connection to all FIA remote offices and City A's datacenter. The total bandwidth to the Service Provider's CPE is 40Mbps. A DMZ (Demilitarized Zone) zone is created to provide secure Internet services. The DMZ is separated from Campus network through CheckPoint firewall. To enhance the security, one Cisco PIX firewall is deployed to separate the external network and the DMZ zone. Inside the DMZ, one layer 2 switch is deployed. The layer 2 switch provides the network connection to the servers in DMZ.

In the Campus Core network, there is one cluster core switch. It is clustered by two physical switches using Cisco's stack technology. This core switch behaves like a Layer 3 switch to provide routing and switching for City B's datacenter. There is also a Top-of-Rack switch for datacenter servers' network connections.

In the Campus Access network, there are three switches which form a ring topology to provide desktop, laptop and printers access in City B's offices.

3.1.2 New City B Network Topology

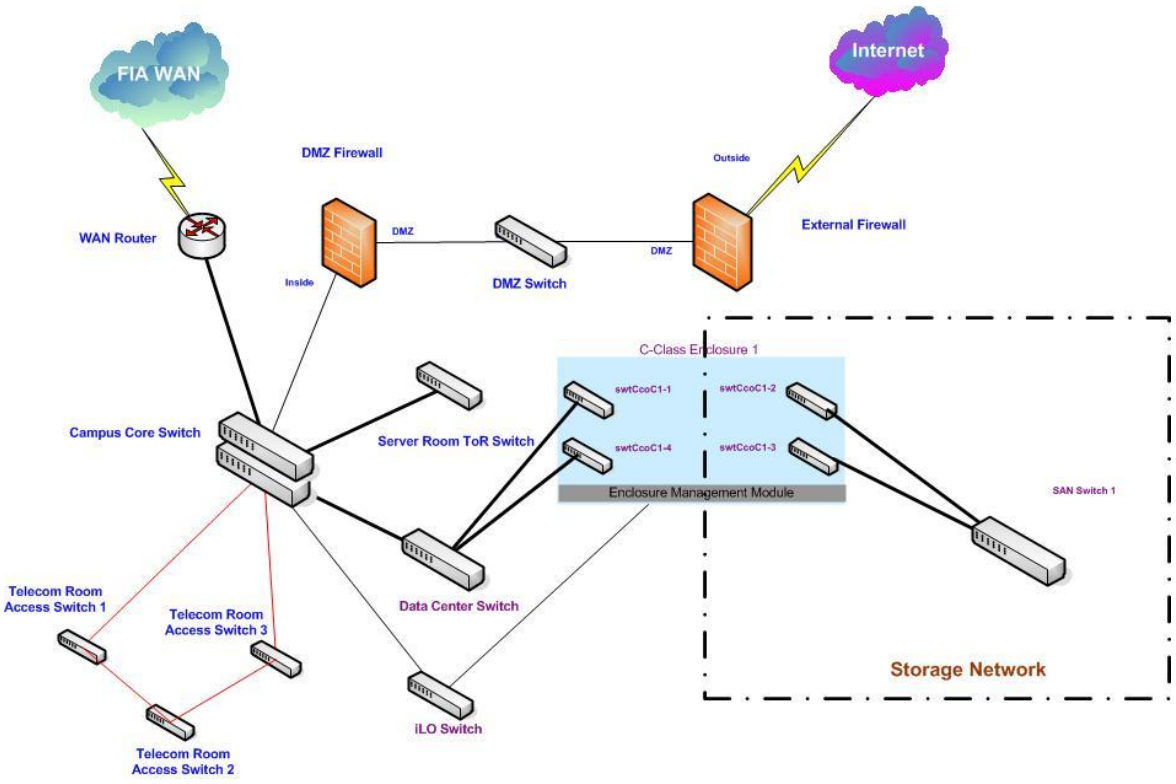

Figure 3-2: City B New Network Logical Topology

To fulfill the network requirements of all business and systems for DRP phase 1, several changes are to be made to the network infrastructure. The changes focus on three areas: storage local area network, datacenter switch, management network and DMZ switch.

Storage local area network is a critical part for the DRP site. The RPO time target relies heavily on the storage local area network. It provides flexible, safe and expandable storage spaces for all IT systems within FIA. In the new storage local area network, the clients implement the iSCSI protocol. This protocol is based on Internet Protocol (IP). All SCSI traffic is segmented and encapsulated in normal IP packets on the client devices. On the storage devices, these packets are de-encapsulated and reassembled back into SCSI traffic. Because the clients and storage reside in the same subnet, the storage network is layer 2 only. For security reasons, this network is isolated from the normal data network. No traffic is to be transferred over other networks. One Cisco Catalyst 4948E switch is selected as the core switch for the storage network.

Datacenter switch provides network access to all DRP systems. In the new DRP network, a Cisco Catalyst 4948E switch is deployed as a datacenter switch. It provides higher port density with up to 48 Gigabit Ethernet ports. It also has higher backbone bandwidth and packets forwarding speed. Separation of datacenter switch and campus core switch offloads traffic from the campus core switch. It also provides the capability to deploy security control between the campus and the datacenter.

The site in City B is a remote DRP site without assigned technical professionals. Remote management is required to perform routine maintenance such as shutdowns or reboots of devices. Currently, all servers have out of band management network interface such as iLO (Integrated Lights-Out). It provides additional management access even when the server is down. To provide the network connections for those management interfaces, one Cisco Catalyst 3560 switch with 48 ports is provisioned.

DMZ is the zone facing the Internet through external firewall. It also connects to internal network through a DMZ firewall. There is a Cisco switch inside DMZ. It provides network connections for all of the devices in DMZ. Because it is only a layer 2 device, it is not scalable to implement multiple VLANs in DMZ. However, there are more services which require multiple network cards with different subnets. Layer 3 capable switch is able to address these requirements. In the new network infrastructure, one Cisco Catalyst 3750 switch with layer-3 feature is used to replace the existing switch.

3.1.3 Storage Network Topology

In FIA DRP phase I, storage network provides network connections for local DRP servers to storage devices. There is no data replication between City A's storage network and City B's storage network. All traffic is in the local network. The protocol between server and storage devices is iSCSI (Internet Small Computer System Interface). All the data packets are encapsulated with IP and switched through Ethernet network.

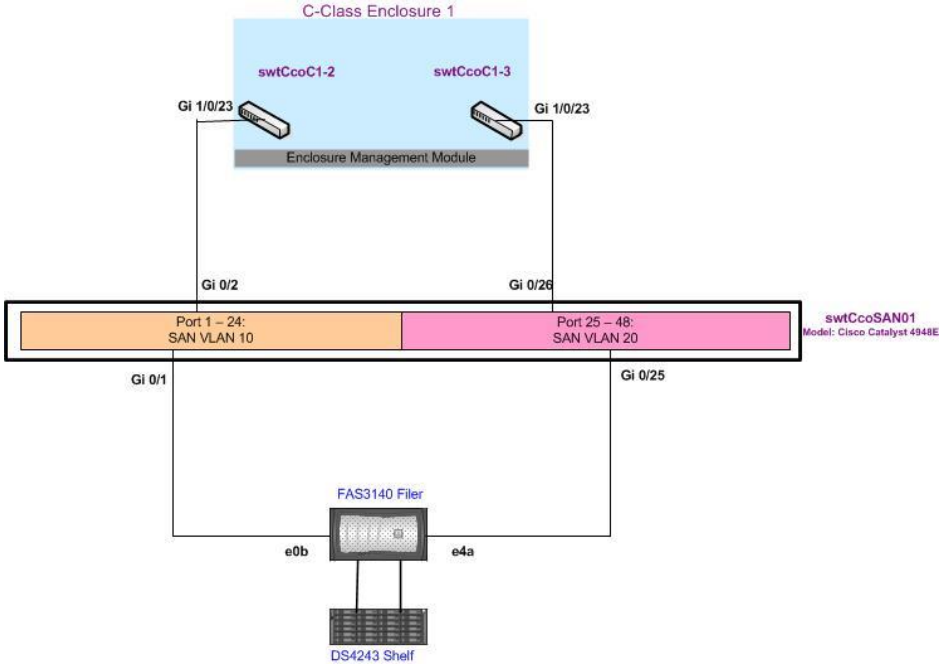

Figure 3-3: City B Storage Network Logical Topology

Figure 3-3 shows the logical topology of City B's storage network. The core network device is a Catalyst 4948E switch. It has two VLANs: 10 and 20. These two VLANs are not routable. The storage device sits behind a NetApp FAS3140 filer. This filer is attached directly to a disk array. For network connection, the filer has two network cards.

One connects to VLAN 10 and the other connects to VLAN 20. On the server side, blade servers that reside in the HP C-Class Enclosure need network storage services.

To support failover, each blade server has two connections to the storage network. One is for VLAN 10 and the other is for VLAN 20. The HP C-Class enclosure provisions these two connections through two blade switches which are installed on the back of the enclosure. Each switch provides one network connection to each blade servers through internal 1 Gbps downlink. For uplink connections, each blade switch has one connection to the external Catalyst 4948E device. The difference between these two blade switches is that one connects to VLAN 10 and the other connects to VLAN 20.

3.2 Exchange Replication Network Consideration

FIA's mail system is built on the Microsoft Exchange 2010 environment. MS Exchange 2010 system requires dedicated replication network to replicate the mailbox databases between nodes within one DAG (Database Availability Group).

There are specific network performance requirements for replication network which are requested by FIA Exchange team:

- Bandwidth: The average bandwidth should be over 8 Mbps and the peak bandwidth should be up to 16 Mbps between any two members within the DAG.
- Round-Trip Latency: The round-trip time should be less than 400 milliseconds between any two members of the DAG.

3.2.1 The Replication Network Topology

In FIA DRP phase I, there is no additional WAN link between City A and City B's datacenters because of business and resource restrictions. The current Service Provider's WAN link between City B and City A is also utilized by the Exchange DAG for the replication link.

Figure 3-4 demonstrates the Exchange replication network's logical topology.

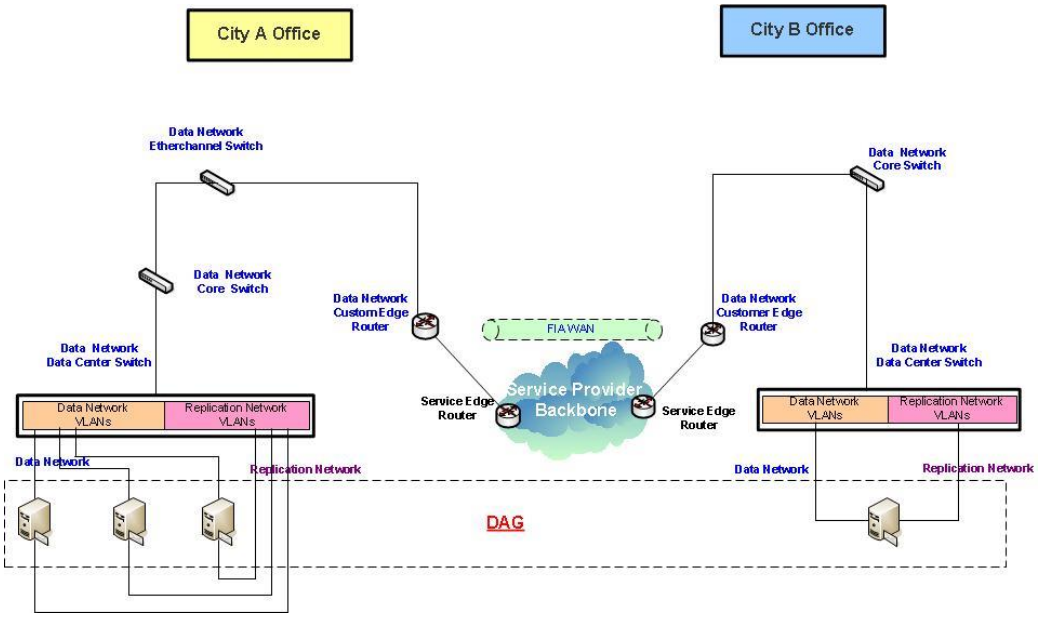

Figure 3-4: Replication Network Logical Topology

As shown in the diagram, there are specific replication VLANs created in City A and City B's datacenters. These VLANs are added to the routing table and can reach each other through the data network. In each node (server) of the DAG, there are two NICs (Network Interface Cards). One is for the data network, and the other is for the replication network. To route the replication traffic through the replication network cards between City A and City B, one static route through replication network card is added into the routing table of each node. So each node will send the synchronization traffic through the replication network card to the rest of the nodes within the same DAG.

3.2.2 Bandwidth Utilization Analysis

Monitoring System

FIA's Exchange DRP solution relies heavily on the replication network's performance. A decrease in available bandwidth between nodes in one DAG will increase the time to synchronize the database. Consequently, the longer synchronization time can cause more data loss when disaster occurs. The RPO time is increased as a result.

A review on the end to end path between Exchange servers in City A and Exchange servers in City B was conducted. The review identified the WAN link between City A and City B to be a potential bottleneck. It has the minimum available bandwidth of 40 Mbps in the end to end path. Does this WAN link meet the FIA exchange system requirements on bandwidth and round-trip delay? In order to answer this question, we set up a monitoring system to gather traffic pattern of the WAN Link and analysis the link usages.

Two monitoring systems were implemented to monitor the WAN link bandwidth usage. One is the Cisco Network Analysis Module which is based on the NetFlow protocol. This monitor tool resides in City A's core switch as a module. The edge router enables the Netflow feature and collects the IP traffic statistics. It sends the collected IP traffic

statistics to the Cisco Network Analysis Module. Network Analysis Module then generates the report of throughput usage of the WAN interface.

The other tool we use is the Opnet ACE Live appliance (The name is changed to AppResponse Xpert). This ACE Live appliance is attached to City A's Etherchannel switch using SPAN (Switched Port Analyzer) port connection. The Etherchannel switch copies each packet of the Edge router to ACE Live appliance. The ACE Live appliance stores and analyzes the IP and TCP headers of the SPANned packets. With these data, it can monitor the bandwidth usage of the WAN link in real time.

The architecture of the monitoring system is illustrated in Figure 3-5 :

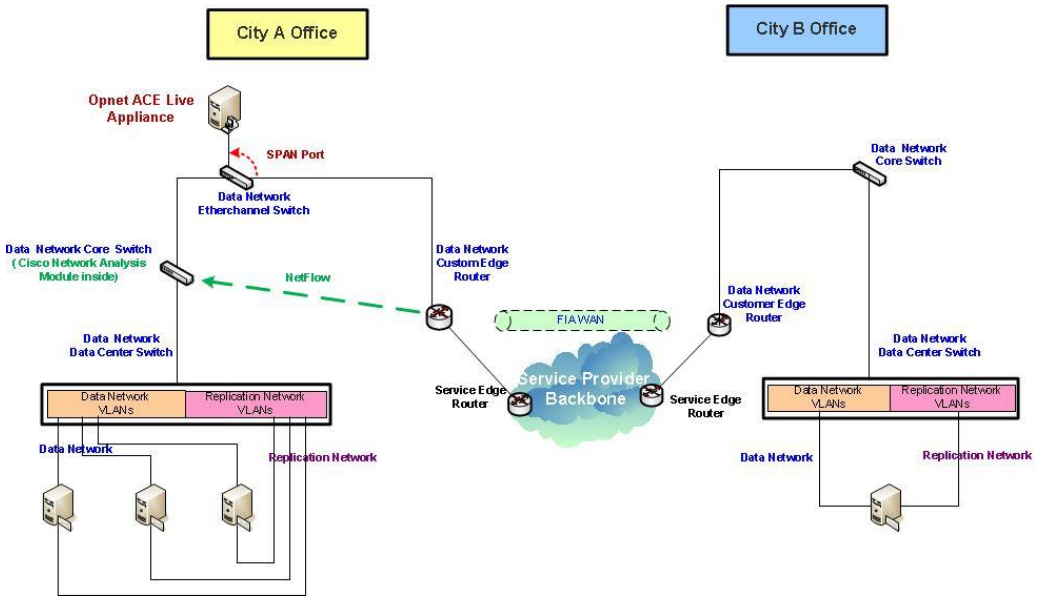

Figure 3-5: Bandwidth Monitor Tools

Bandwidth Utilization

Based on one month's data, the peak bandwidth usage of the WAN link between City A and City B is less than 20 Mbps (hourly average).

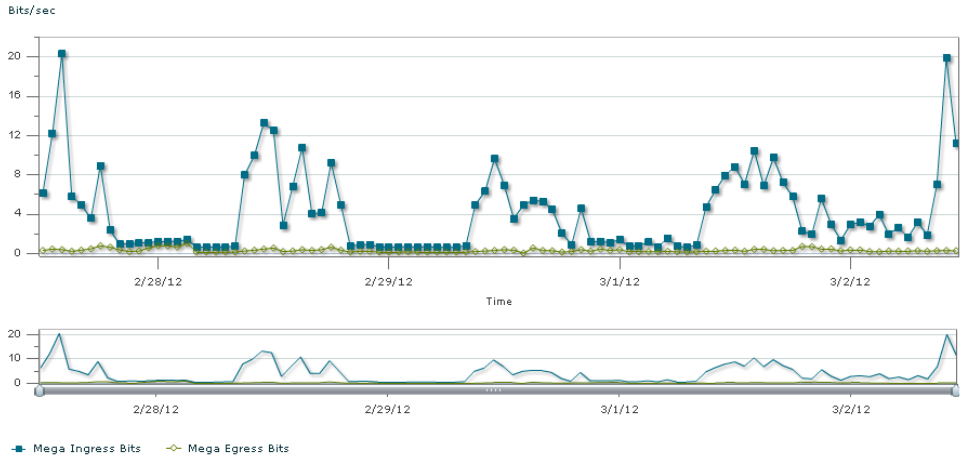

Figure 3-6: WAN Link Traffic Usage

Figure 3-6 shows the outbound traffic usage of City A's router WAN interface of a week's span. The maximum bandwidth usage from City A to City B is 20 Mbps in this period. Considering the worst case in which the peak of FIA exchange replication traffic (16 Mbps) also occurs at the same time, the total required bandwidth would be 36 Mbps (20Mbps + 16Mbps). If we compare this required bandwidth (36 Mbps) with the existing WAN bandwidth (40 Mbps), the existing WAN link bandwidth meets the required bandwidth when FIA exchange DRP system is deployed.

The bandwidth usage in the above figure is hourly average bandwidth usage. If we look at the average bandwidth usage per minute, it can reach highest bandwidth and causes network congestion. As a result, replication traffic is not forwarded, and is queued in Exchange servers during the network congestion. Once the bandwidth is available, the replication traffic will be forwarded again.

3.2.3 Round-trip Delay Analysis

Sample Test Using ICMP ping

Neither of our monitoring tools, Cisco Network Analysis Module and Opnet ACE Live appliance, provides the network round-trip time between City A and City B. To obtain information of the round trip time, the ICMP (Internet Control Message Protocol) ping tool is utilized.

In this exercise, one Windows machine in City A datacenter is used as the test source machine. The destination device is on City B's Campus Core switch. Tester executes the following command:

ping xxx.xx.0.1 –l 1500 –n 100

The above command sends 100 ICMP request packets and each packet is 1500 bytes.

During one week, we ran the test five times per day. All the tests are executed during business hours between 8:15am and 4:30pm when testers are available.

The test results show the maximum network round trip time is not over 50ms. This is far below the required 400ms.

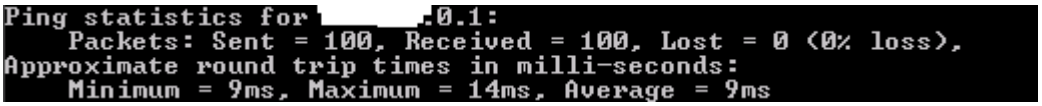

Figure 3-7: Round Trip Test Result Sample

Note: The first two numbers of IP address are hidden for confidentiality.

The above screenshot is one of the test results. It demonstrates that the maximum round trip return time is 14ms with an average of 9ms.

Continuous Monitoring the Round-trip Using Cisco IP SLA

This sample test has obvious limitations. It does not test the round-trip time in all possible network states. For example, what is the round-trip time when the WAN bandwidth utilization is very high? To obtain the network round-trip time between City A and City B's replication network, a monitoring system has been set up using the Cisco IP SLA

(Service Level Agreement) feature. It monitors the round-trip time of the network for one week.

Cisco IP SLA is a feature provided with Cisco's routers and switches. It uses generated traffic to measure network performance between two networking devices. It generates specific traffic from one Cisco device to another. Based on the returned traffic, IP SLA calculates the network characteristics between these two devices, such as round-trip time, packet loss and so on. IP SLA stores these data on the devices with a limit buffer. An external network monitor tool which supports Cisco IP SLA can query these IP SLA data through SNMP poll from the Cisco device. Then the network monitoring tool presents those data in tables or graphs for analysis.

Cisco IP SLA can generate several different types of traffic to measure the metrics of the network. In FIA DRP phase I, three different types of traffic are generated to measure the round-trip time over the replication network between City A and City B: ICMP (Internet Control Message Protocol) Echo, TCP (Transmission Control Protocol) Connect and UDP (User Datagram Protocol) Echo. One IP SLA capable network monitor tool is implemented. This monitor tool is called PRTG Network Monitor tool which is the product of Paessler AG Company.

This monitoring architecture is demonstrated in Figure 3-8:

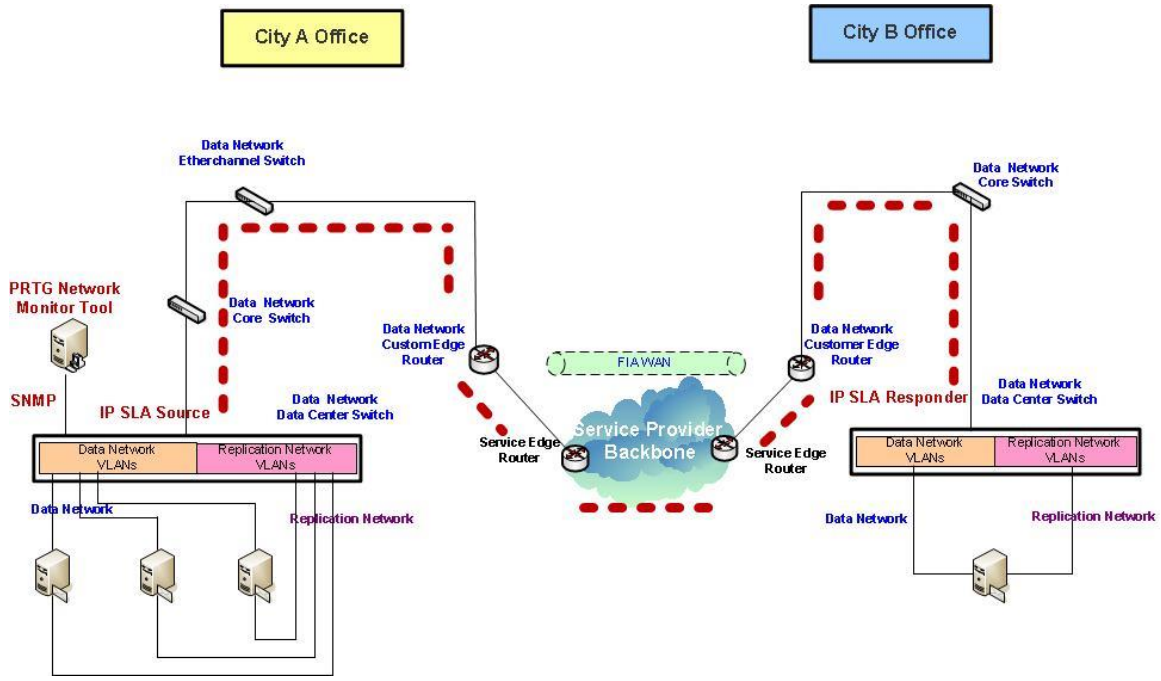

Figure 3-8: IP SLA Monitor Architecture

Figure 3-8 shows how IP SLA's monitoring system is implemented. The Data Center switch in City A is configured as the IP SLA source. The IP SLA responder is configured at the Data Center switch in City B. The source device sends a generated packet to the responder device. After the responder device receives the packet, depending on the type of the IP SLAs operation, it responds with time-stamp information for the source to make the calculation on round-trip time.

Three types of IP SLAs are configured to measure the round-trip time: ICMP echo, TCP Connect and UDP Echo. The interval at which each type of IP SLAs is repeated is 30 seconds. The size of the ICMP request packets and UDP packets is 1200 bytes. This is the size of the Exchange replication traffic packets.

The PRTG Monitor Tools is configured to query the IP SLA data in IP SLA source device every 60 seconds. It stores the data in a local database and generates reports for each IP SLAs at 0:00 on a daily basis.

The whole monitoring activity continued for a week. Based on the reports, the maximum round-trip time between the IP SLA source and responder device is 56 milliseconds which is shown in the report of TCP Connect IP SLA. The samples of the reports are in the Figure 3-9, Figure 3-10 and Figure 3-11. The Figure 3-12 shows the traffic throughput of the City A's WAN link in the same period as the Figure 3-9, 3-10 and 3-11.

Report Time Span:	7/12/2012 12:00:00 AM - 7/13/2012 12:00:00 AM				
Report Hours:	24 / 7				
Sensor Type:	Cisco IP SLA (60 s Interval)				
Probe, Group, Device:	Local probe > 1st group > datacenter				
Uptime Stats:	Up:	100 % [23h50m16s]	Down:	0 % [0s]	
Request Stats:	Good:	100 % [1432]	Failed:	0 % [0]	
Average (Average Round Trip Time (RTT)):	8 msec				

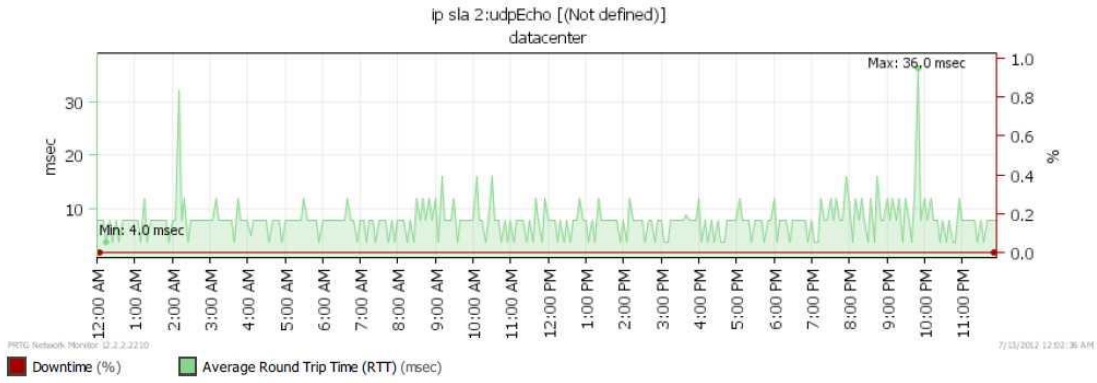

Figure 3-9: IP SLA UDP Echo Round Trip Report

Figure 3-9 demonstrates the round trip time in a one day period based on IP SLA UDP Echo. The average round trip time of one minute within a day is less than 15 milliseconds. The maximum round trip time is 32 milliseconds.

Report Time Span:	7/12/2012 12:00:00 AM - 7/13/2012 12:00:00 AM				
Report Hours:	24 / 7				
Sensor Type:	Cisco IP SLA (60 s Interval)				
Probe, Group, Device:	Local probe > 1st group > datacenter				
Uptime Stats:	Up:	100 % [23h50m6s]	Down:	0 % [0s]	
Request Stats:	Good:	100 % [1432]	Failed:	0 % [0]	
Average (Average Round Trip Time (RTT)):	11 msec				

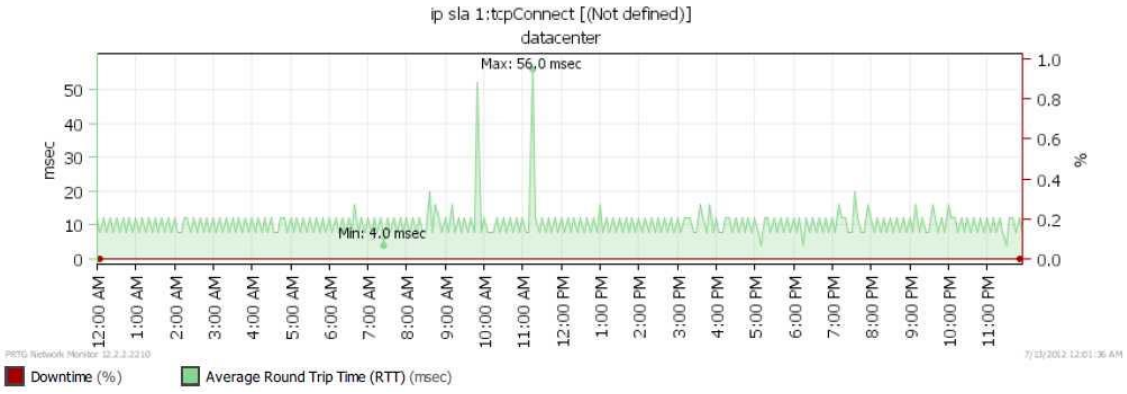

Figure 3-10: IP SLA TCP Connect Round Trip Report

Figure 3-10 demonstrates the round trip time in a one day period based on IP SLA TCP Connect. The average round trip time of one minute within a day is less than 20 milliseconds. The maximum round trip time is 56 milliseconds.

Report Time Span:	7/12/2012 12:00:00 AM - 7/13/2012 12:00:00 AM				
Report Hours:	24 / 7				
Sensor Type:	Cisco IP SLA (60 s Interval)				
Probe, Group, Device:	Local probe > 1st group > datacenter				
Uptime Stats:	Up:	100 % [23h49m0s]	Down:	0 % [0s]	
Request Stats:	Good:	100 % [1431]	Failed:	0 % [0]	
Average (Average Round Trip Time (RTT)):	8 msec				

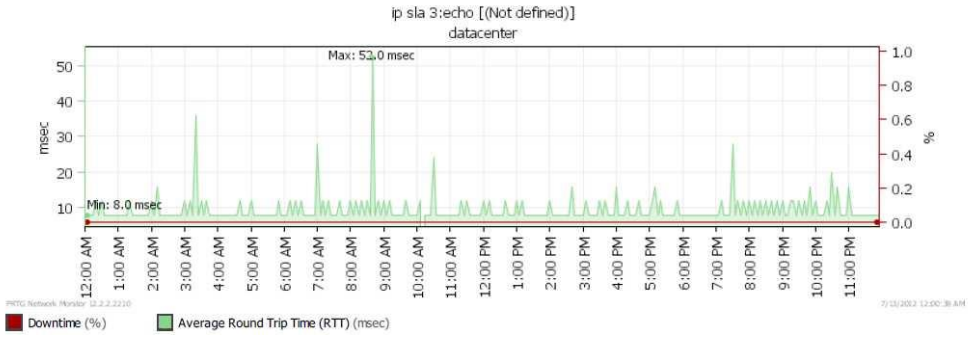

Figure 3-11: IP SLA ICMP Echo Round Trip Report

Figure 3-11 demonstrates the round trip time in a one day period based on IP SLA ICMP Echo. The average round trip time of one minute within a day is less than 20 milliseconds. The maximum round trip time is 52 milliseconds.

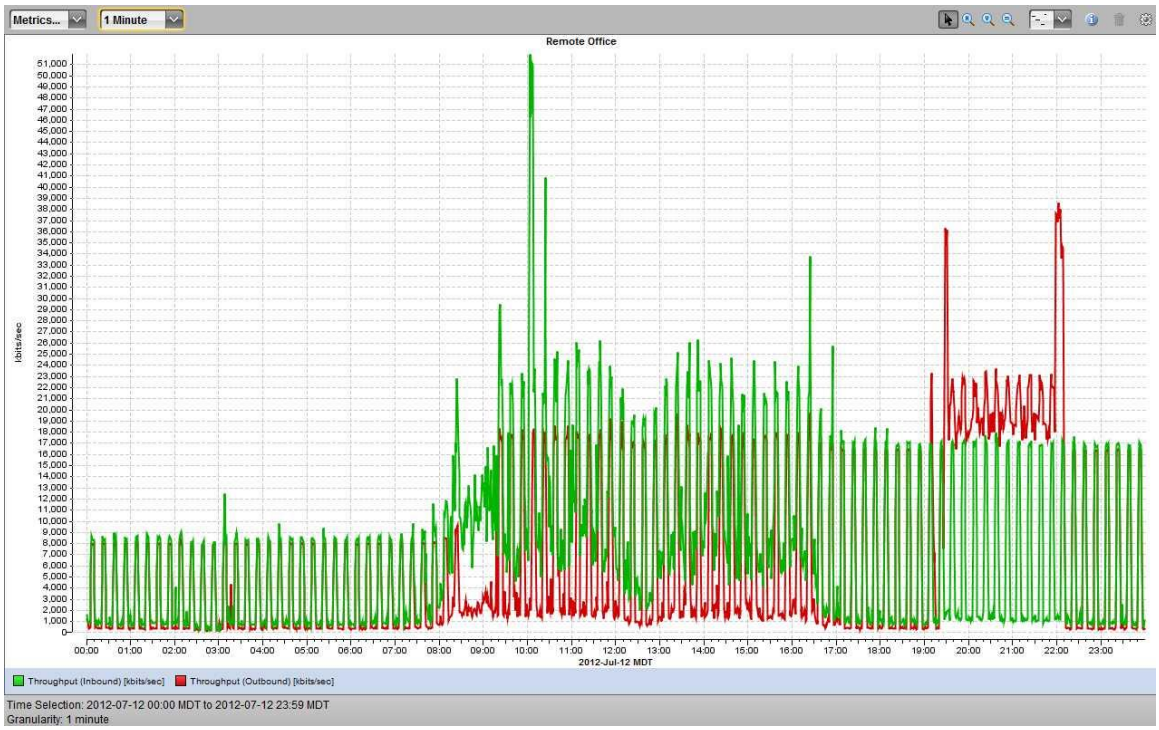

Figure 3-12: Throughput of City A's WAN Link

Figure 3-12 demonstrates the throughput of City A's WAN link in the same day as Figure 3-8, 3-9 and 3-10. This report is generated by Opnet ACE-Live monitor tool. It records the inbound and outbound throughput every minute. The inbound traffic is from City A's data center to the remote office. The outbound traffic is from the remote office to City A's data center. The peak traffic throughput is up to 52 Mbps at 10:00am. (Note: This test was conducted after the WAN link bandwidth was upgraded from 40 Mbps to 60 Mbps.)

3.3 IP Addresses and VLANs

In DRP phase I, significant modifications are required in City B's datacenter network. IP address allocation and VLAN plan are two major components that need to be modified. Two objectives are identified for IP address allocation. The first is scalability. The new design should be able to accommodate future requirements such as new DRP services for next DRP phase. The other target is to shrink the routing table in the router. To achieve these two targets, the VLANs and IP subnet information is listed as below:

Category	IP Address Range	Usage	VLAN
Internet	xxx.xx..113.0–xxx.xx.113.255	City B Internet Public IPs	
WAN	xxx.xx.112.0– xxx.xx.112.255	Service Provider Interconnection Network	
	xxx.xx.1.0 – xxx.xx.1.255	FIA WAN Network	
Data Network LAN	xxx.xx.0.0 – xxx.xx.0.255	City B Network Edge	2
	xxx.xx.1.0 – xxx.xx.1.255	City B DMZ	1
	xxx.xx.5.0 – xxx.xx.5.255	City B ILO Management	5
	xxx.xx.6.0 – xxx.xx.6.255	City B Server Data Connection; Vmware Host Service Console; Vmware Host; Vmware Guest Data Connection	6
	xxx.xx.8.0 – xxx.xx.8.255	City B Office Desktop, Printer	8
	xxx.xx.23.0 – xxx.xx.23.255	City B legacy server (DHCP server)	1
Heart Beat	yyy.yy.1.0 – yyy.yy.1.255	Vmware Host Heart Beat Network	901
Replication Network	xxx.xx.96.0- xxx.xx.96.255	Exchange Replication Network	900
SAN LAN	yyy.yy.10.0 –yyy.yy.10.255	SAN Network 1	10
	yyy.yy.20.0 – yyy.yy.20.255	SAN Network 2	20

Note: For confidentiality, the first two octets of the IP address are hidden.

3.4 Routing Design

In FIA's internal network, Enhanced Interior Gateway Routing Protocol (EIGRP) is the dynamic routing protocol which is implemented across the whole corporation network. Although EIGRP is Cisco proprietary routing protocol, it still meets the current FIA requirement.

3.4.1 EIGRP Architecture

FIA EIGRP routing architecture is a hub-spoke topology. City A and City B's WAN (Wide Area Network) edge routers are the hub routers. Each remote office's WAN edge router is the spoke router. The spoke router has two logic network connections. One connection is to City A's datacenter and the other one is to City B's datacenter. The two Hub routers are connected to each other through direct logical connections.

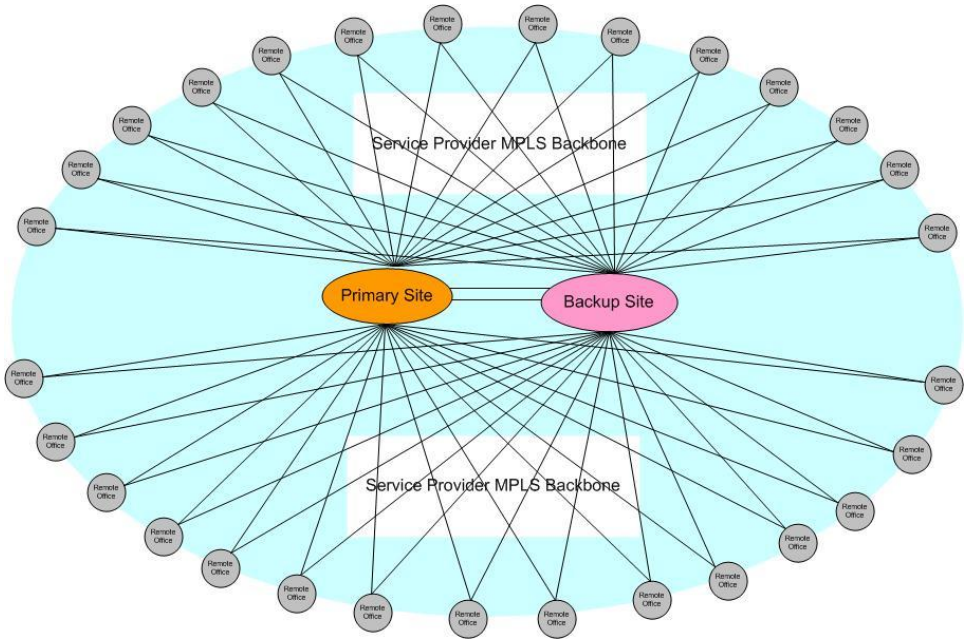

Figure 3-13: FIA WAN Network Topology

Figure 3-13 shows the high-level logical WAN network topology of FIA. It demonstrates the logical network connection of the WAN network. City A's datacenter (Primary site) has one WAN edge router named rtrCityA. The WAN edge router in City B's datacenter (Backup site) is named rtrCityB. Each branch office edge router connects to both rtrCityA and rtrCityB.

3.4.2 EIGRP Topology Table

The central algorithm of EIGRP is the Diffusing Update Algorithm (DUAL) which selects the optimum route to destination networks for the routers in the network based on metrics. To compute the metrics, EIGRP DUAL relies on data structure in neighbor table and topology table which are built based on hello protocol and reliable transport protocol.

The neighbor table includes a list of adjacent routers which are running EIGRP process. The EIGRP topology table includes all the information received from EIGRP neighbors as well as information received from other routing protocols such as static route.

There are two types of route entries in EIGRP topology table: internal EIGRP route and external EIGRP route. The internal EIGRP routes are the routes within the same autonomous system (AS). Since all of FIA's internal networks are under FIA corporate control, all subnets of FIA's internal networks are on the internal EIGRP routes in the EIGRP topology table. The external EIGRP routes include the routes which are redistributed from other sources which are not in the same autonomous system. The default route to access Internet is redistributed in to FIA EIGRP AS through static route. In the FIA EIGRP topology table, this default route is external EIGRP route.

Inside the EIGRP topology table, each route entry has two layers of metrics: vector metrics and composite metric. Vector metric is a six-element vector containing parameters (bandwidth, delay, load, reliability, hop count, and MTU) that describe the distance between a router and the destination subnet. The vector metric is used in all EIGRP routing updates. In the FIA EIGRP topology table, those parameters of the vector metrics are calculated by the EIGRP DUAL based on the default setting on the router.

The default vector metrics of internal routes are computed based on link or interface information:

Bandwidth (BW)	*BW of interface*
Delay (DLY)	*Delay received +Delay of interface*
Load	*Load of interface*
Hop Count	*HopCount received +1*
MTU (Maximum Transport Unit)	*MTU of interface*
Reliability	*255*

The default vector metrics of external routes are listed as below:

Bandwidth	*1000000 kbps*
Delay (DLY)	*10 microseconds*
Load	*1*
Hop Count	*1*
MTU (Maximum Transport Unit)	*1500*
Reliability	*255*

Composite metric is an integer used to compare different routes toward the same destination subnet. It is only used internally in the router and is never sent to EIGRP neighbors. The formula below transforms the vector metric into a composite metric:

*Composite metric = [K1*BW + K2*BW/ (256-load) + K3*DLY] * K5 / (reliability + K4)*

Note: BW = 10Gbps / Bandwidth

In the FIA EIGRP configuration setting, the K-values are kept as the default values. The default values of K1 and K3 are 1, and all the other factors have a default value of 0. The default composite metric is therefore a sum of the total delay and the inverse bandwidth.

When there are multiple paths to the same destination, EIGRP DUAL will compute the composite metric for each route and store them in the topology table. The route with the

lowest metric is the optimum route to the destination network for EIGRP protocol. EIGRP will provide this route information to the router for insertion to the routing table.

3.4.3 Routing Table

When a layer 3 packet arrives at a router, the router will look up the routing table to find the next hop and forward this packet. Based on Cisco routers' forwarding policy, the route with the longest prefix length that matches the packet's destination network is selected for packet forwarding.

In the FIA network, the router builds the routing table from three different sources: directly connected network, static route and EIGRP route. When there are two or more different routes to the same destination from these three sources, FIA's Cisco router selects the best route based on the lowest administrative distance value. Administrative distance is a measure of the trustworthiness of the source of the routing information. Administrative distance has only local significance, and is not advertised in routing updates. Those sources in FIA routers are kept using the default value as below:

Route Source	***Default Distance Values***
Connected Interface	0
Static Route	1
Internal EIGRP	90
External EIGRP	170

The external EIGRP route has higher administrative distance value than internal EIGRP route.

3.4.4 Disaster Recovery Process in DRP Phase I

When disaster happens in City A, the logical links of each remote office to City A edge router are down. The services will be failed over to use the disaster site in City B. As the routes to City B network are EIGRP internal routes which are already in the routing table, there should be no routing outage for this change.

However, there are considerations for Internet access. There are dual Internet access links in FIA network. City A and City B's data centers both have independent Internet access links. The business policy requires all Internet traffic to only go through City A's Internet access link. The only condition in which the Internet traffic goes through City B's Internet access link is when City A is not available.

The traffic to the Internet is managed by the default route on the remote office routers. To achieve the Internet access policy, City A's edge router redistributes a static default route which points to City A's Internet link into EIGRP topology table and propagates it to all remote office routers during the normal operation stage. When disaster happens in City A, the City B Edge router needs to generate default route to remote office routers and direct traffic through City B's Internet access link. To implement this change in the EIGRP routing protocols, there are two methodologies: manual update and automatic update.

The manual update method requires the network administrator of DRP team to manually add a new static default route which points to City B's Internet link. EIGRP redistributes it into the EIGRP routing protocol in City B's edge router after disaster occurs. Then City B's edge router advertises this external EIGRP route to all remote office routers.

The automatic update method requires the default route in remote office routers to be automatically changed to City B's edge router and access Internet through City B's Internet link. This automatic update is controlled by the EIGRP metrics.

Comparing these two methodologies, the automatic update is a more efficient solution than manual update. The automatic changes can be carried out in seconds. As a result, the RTO will be in seconds too. The only drawback is that FIA network team is not familiar with this EIGRP automatic process. To complete the project on time, FIA network team decided to adopt the manual change method in DRP phase 1. The detailed execution plan is captured in Section 4.2.1 of this document.

3.4.5 Disaster Recovery Process Improvement in Next DRP Phase

For Internet access, automatic update of default route in remote office routers through EIGRP protocol is recommended. To help FIA network team to understand and validate the automatic process, a similar environment was built in the FIA Lab.

Lab Topology

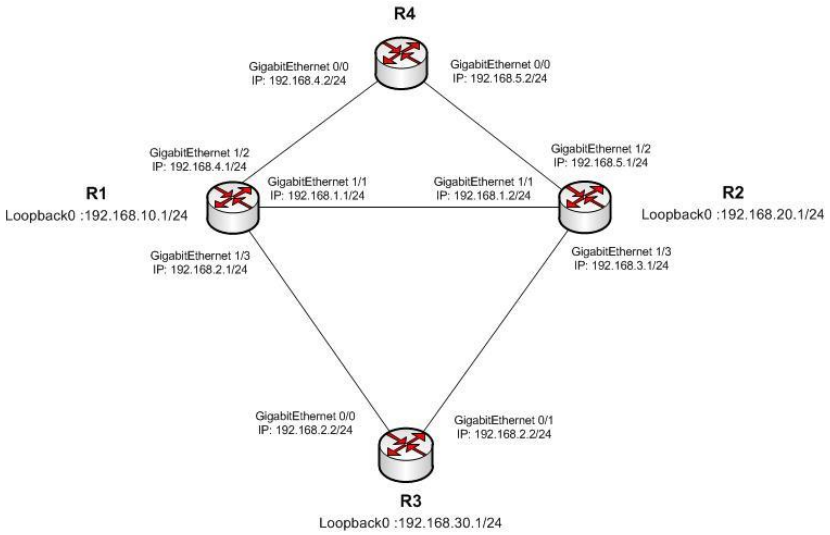

Figure 3-14: EIGRP Lab Topology

The above diagram demonstrates the lab topology. Each device simulates the routing function of the FIA network devices. R1 acts as the edge router in City A. R2 simulates the edge router in City B. R3 simulates edge router in one remote office. R4 behaves as Internet access.

R1 has three 1 Gbps Ethernet links. One link to R4 is as Internet link. One link to R2 simulates the inter-connect link between City A and City B. The last link to R3 is the access link for remote offices. For R2, it has three links similar to that of R1. One of the differences is that R2 has its own independent Internet link.

Configuration Example

In this Lab exercise, the following functions need to be achieved automatically by EIGRP routing protocol:

- R1's default route always points to R4
- When R1 is alive, R2 and R3's default route should point to R1
- When R1 is not alive, R2's default route should point to the R4 and R3's default route should point to R2.
- When the link between R1 and R3 fails, R3 should change the default route from R1 to R2. R2's default route should remain to be R1
- When the link between R1 and R2 fails, R2 should change the default route from R1 to R3. R3's default route should remain to be R1

The relevant portions of the R1 and R2 router configuration are shown in Figure 3-15:

```
hostname R1
!
router eigrp 2
 router-id 192.168.10.1
 network 192.168.1.0
 network 192.168.2.0
 network 192.168.4.0
 network 192.168.10.0
 redistribute static metric 10000000 10 255 1 1500
!
! Default route toward the R4
ip route 0.0.0.0 0.0.0.0 192.168.4.2
```

```
hostname R2
!
router eigrp 2
 router-id 192.168.20.1
 network 192.168.1.0
 network 192.168.3.0
 network 192.168.5.0
 network 192.168.20.0
 redistribute static metric 1000000 10 255 1 1500
!
! Default route toward the R4
ip route 0.0.0.0 0.0.0.0 192.168.5.2 200
```

Figure 3-15: Sample Configuration of R1 and R2

R1 and R2 both have configuration to redistribute default route into the EIGRP topology table. However, there are two different configuration settings in R2 comparing to R1. These two changes guarantee that R2's static route is not effective when R1 is alive.

One difference is that the administrative distance of "ip route 0.0.0.0" is increased from 1 to 200. This change is to have this route not inserted into R2's routing table when R1 is alive. In R1's configuration, one default route is redistributed into the EIGRP topology table as an external route. This external route will be sent to R2 through EIGRP protocol. When R2 receives this route, it will compare the administrative distance between this external EIGRP route and its own static route "ip route 0.0.0.0". The lower value of the administrative distance is selected. Because the external EIGRP route's administrative distance is 170, it is lower than the 200 of the R2's local static route "ip route 0.0.0.0". The external EIGRP route will be inserted into the R2's routing table. So that, R2's default route points to R1. As a result, R2 doesn't redistribute its local "ip route 0.0.0.0" into the EIGRP topology table. The R2's EIGRP topology table only has one default route which is external EIGRP route from R1.

Another difference is that the bandwidth of "redistribute static metric" in R2 is decreased from 10000000 to 100000. This change is to increase the composite metric when R2 redistributes its local static default route into EIGRP topology table. When EIGRP process compares these two default routes from R1 and R2, the route from R1 will take

effect as the lower composite metric value. It avoids default route change in remote offices under some uncertain situations where R2's local static default route takes effect.

The relevant portions of R3 router configuration are shown in Figure 3-16:

```
hostname R3
!
router eigrp 2
 router-id 192.168.30.1
 network 192.168.1.0
 network 192.168.3.0
 network 192.168.30.0
 !
```

Figure 3-16: Sample Configuration of R3

Validation

In this Lab environment, the default route in R1, R2 and R3 has been examined in three failure scenarios.

Scenario 1: Link between R1 and R3 fails. In this scenario, the R3's default route is changed to R2. The R2's default route still points to R1.

Scenario 2: Link between R2 and R1 fails. In this scenario, R3's default route still point to R1. R2's default route is changed from R1 to R3.

Scenario 3: R1 fails. In this scenario, R2's default route is changed to R4. R3's default route is also changed from R1 to R2.

In conclusion, the test result shows the automatic default route update relying on EIGRP protocol is practical, and it satisfies FIA's Internet traffic policy requirements.

3.5 Network Service DRP Design

3.5.1 DHCP DRP

DHCP service is one of the basic network services within the network. It automatically allocates IP addresses to end devices such as desktop and laptop with network parameters. In FIA's network, DHCP service is built for DRP requirement. There are two DHCP servers. One is located in City A's datacenter. The other is located in City B's datacenter. Both servers are online. When the end device sends DHCP request to a DHCP server, that DHCP server will assign one IP from its IP address range to the end device.

The difference between the two DHCP servers is the number of available IP addresses. Based on office sizes, City A's DHCP server hosts IP addresses range from xxx.xxx.xxx.64 to xxx.xxx.xxx.209. It has in total 146 available IP addresses. City B's DHCP server hosts IP addresses range from xxx.xxx.xxx.210 to xxx.xxx.xxx.254. It has in total 45 IP addresses. When City A's DHCP server fails, the City B DHCP server has enough IP addresses to meet the business requirement.

Although there are two DHCP servers in both datacenters, correctly forwarding the DHCP request packets is a critical task. The DHCP request packets are broadcast packets, and the end device may not be in the same subnet as the two DHCP servers. The solution relies on the DHCP relay feature of the Cisco router. The Cisco router is configured as two ip-helper servers. One points to City A's DHCP server and the other points to City B's DHCP server. When the router sees a DHCP request broadcast packets, it redirects the packet to the two DHCP servers using unicast packets. When the router receives the response packets, it will forward them to the end devices. The end device will only accept the first arriving DHCP response packet. The end device will ignore the other DHCP response packet.

3.5.2 DNS DRP

Domain name system (DNS) provides mapping between domain name and IP address. Domain name is convenient for users. However, computers use IP address to communicate with other devices. Numeric address is effective for computer processing. DNS works as a directory lookup service to provide translation between IP address and the domain name. It can map the domain name to IP address, and vice versa.

FIA's DNS system can be classified to two DNS systems. One is Internal DNS system. And the other is External DNS system.

3.5.2.1 Internal DNS System

FIA Internal DNS Infrastructure

Internal DNS system provides name resolution service for internal devices. It hosts internal domain names. FIA internal name servers use Service Provider DNS servers to resolve the domain names which are not hosted or cached locally. The internal DNS servers work recursively and forward to service provider's DNS system to resolve the name.

FIA's Internal DNS system has four name servers. All of the name servers are built on Microsoft Windows 2008 R2. The DNS services are integrated with Microsoft Active Directory. Each name server is also the domain controller which runs Active Directory (AD) service. One identical database is maintained by the four name servers. Each name server gets one copy of this database and stores it in the AD. The changes made on any one of the four name servers will be updated in the AD. Active Directory is responsible for replicating the data to other domain controllers. The DNS services in other domain controllers poll the updates from the Active Directory. Physically, two name servers are located in City A's datacenter with the other two name servers located in City B's datacenter. Based on Active Directory's configuration, the replication interval between the two sites is 15 minutes. It means that the changes made at City A's name servers will take up to 15 minutes to be replicated by City B's name servers.

Resource Records with Scan Service

There are two types of resource records (RRs) in the internal name servers' database. One is static RR and the other is dynamic RR. The static RRs are manually created and maintained by DNS administrator. As those RRs are not changed often, they are quite stable.

The dynamic RRs are created by end devices and scavenged by the timer setting in the name servers. The dynamic RRs allow authenticated end devices register their name with their current IP addresses in the name servers' database. It dramatically offloads the DNS administrator's workload. The end devices which use this dynamic methodology are computers which are operated by the Microsoft Windows operating system.

In FIA production environment, scanning service relies on the dynamic RRs. The scanning service leverages those dynamic RRs to send the scanned documents to the end devices. In the scanning machine, there is one specific user profile for each user. The user's computer's internal domain name is configured as the destination where the scanned documents are sent. Since the user's computer receives IP through DHCP, there is no static IP address allocated for each user's computers. Thus the user computers must dynamically register their domain names with current IP addresses in internal names servers every time after they successful login to the FIA's AD internal domain. With accurate dynamic RRs, scanning machine is able to locate the correct IP address and deliver the documents correctly.

Disaster Recovery Process of Internal DNS system

During the disaster period, the two name servers in City B are responsible to provide the name resolution services for all internal users.

The DNS database replication between City A and City B occurs every 15 minutes. The RTO of internal DNS is zero. And the RPO of internal DNS is 15 minutes. When disaster occurs, only changes on City A's name servers within this 15-minutes window are lost. If changes are made on static RRs, DNS administrators need to update those records on City B's name servers manually. If those changes are altered by end devices on dynamic RRs, such as new computer just renewed its IP address, those end devices need to log off and log in to the domain. Then end devices will automatically update the RR records in City B's name servers.

3.5.2.2 External DNS System

FIA External DNS Infrastructure

FIA's External DNS system provides the FIA.CA domain name resolution for external users. It hosts only the FIA.CA domain zone. The FIA External DNS system has two external DNS name servers. One is in City A's DMZ network and the other is in City B's DMZ network. Both name servers are based on Microsoft DNS servers. City A's name server is the authoritative for FIA.CA zone. The type of zone configured in the server is master. Based on Microsoft DNS service's capability and DNS protocol, only one name server can be acting as master for each zone. Thus all necessary changes to zone FIA.CA can only be made on City A's DMZ name server. The City B's DMZ name server hosts zone FIA.CA as secondary. The secondary has only the read right on the FIA.CA zone. It obtains new zone files from master name server through zone transfers.

The infrastructure and zone transfer of the External DNS system is demonstrated in Figure 3-17:

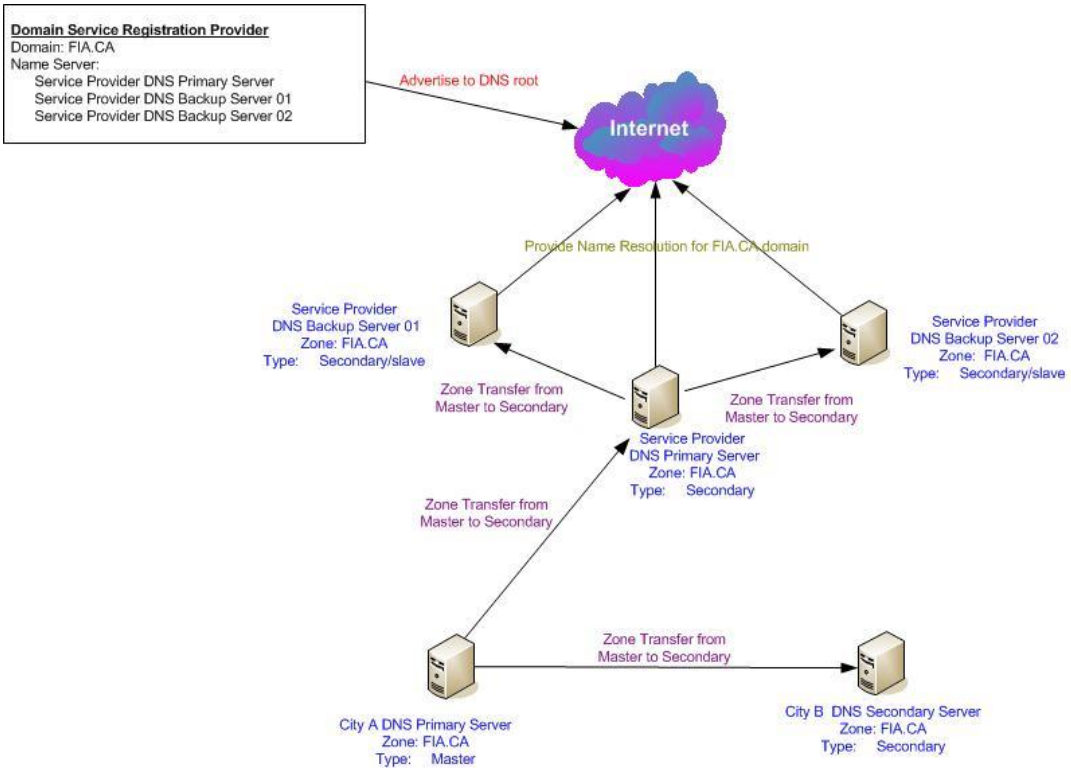

Figure 3-17: FIA External DNS Infrastructure and Zone Transfer

In Figure 3-17, FIA registers the domain FIA.CA in the Domain Service Registration Provider (DSRP). The DSRP advertises three name servers of a Service Provider as the name servers for FIA.CA domain. These three name servers provide the name resolution for FIA.CA domain to Internet users. It offloads the capacity requirement of FIA's name servers, as well as minimizes the security risks from the Internet.

Automatic zone transfer is the mechanism to distribute the zone data from authoritative (master) of the zone to other non-master or secondary name servers. The zone transfer is a hierarchical infrastructure. In the Figure 3-17, the FIA's primary DNS server in City A

hosts the zone of FIA.CA as master. It transfers the zone data to primary name server of service provider and the secondary name server at City B. Then the primary name server of the service provider will act as the master for another two name servers of service provider. The other two name servers receive the FIA.CA zone data from the primary name server of service provider by zone transfer.

There are two methods for a secondary server to initialize the automatic zone transfer: DNS Notify by master and periodical check. Windows-based DNS servers support DNS Notify. When the FIA.CA zone is updated in City A's DNS name server, it sends a DNS notify messages to City B's DNS name server as well as primary DNS name server of the service provider. These two secondary name servers receive the notify message and then respond by initiating a zone transfer request back to City A's DNS name server.

At the same time, secondary DNS servers will also periodically check the changes in the master DNS server. In this type of zone transfer, the SOA (Start of Authority) record in FIA.CA zone determines the check interval and tells the secondary if there are changes in the master name server.

To detect changes, secondary DNS name servers check the SERIAL field of the SOA for FIA.CA zone. Whenever any change is made to FIA.CA zone, the value of SERIAL will be a simple increment. When the value of SERIAL of the SOA in master is higher than the value of SERIAL in secondary, it means there are valid changes in the master server. The secondary server then must request zone transfer from master server. By comparing the SERIAL, it increases the efficiency and improves traffic optimization.

The following parameters in SOA RR define the poll interval:

- **Refresh**— tells the secondary name server how often to poll the primary name server and check for a serial number change. In the SOA of FIA.CA zone, it is 15 minutes. The secondary servers will check the serial number every 15 minutes. The period of 15 minutes is the longest time for the change to propagate.
- **Retry**— the interval at which the secondary name server tries to reconnect with the primary name server, in the event that it fails to connect at the Refresh interval. This interval is 10 minutes in the case of the SOA of the FIA.CA zone
- **Expire**— the interval after which a secondary name server needs to delete the data of the primary name server, if it fails to reconnect to the primary name server. In the SOA of FIA.CA zone, it is 7 days.

Disaster Recovery Process of External DNS

When City A's DNS name server is out of service in disaster period, the three name servers of service provider can still provide the name services for FIA.CA zone to Internet users for seven days. The seven day limit is defined by Expire field of SOA records of the FIA.CA zone.

However, the three name servers of service provider are all secondary for FIA.CA zone. They only cache the RRs of the FIA.CA zone. They cannot make any changes on those RRs. These limitation doses not satisfy the requirement of FIA web server DRP's requirement. In FIA DRP Phase I, the web server in City B acts as standby server. There

should be no traffic to City B's web server during normal operation. It only provides services to Internet client during disaster period. Thus one host resource record in FIA.CA zone maps the web server domain name (www.fia.ca) to the public IP of City A's web server. When disaster happens, this host resource record has to be updated to map to City B's web server public IP.

At the same time, if the City A's master server has not recovered after seven days, all the cached data of FIA.CA zone in name servers of service provider will be discarded.

To resolve this issue, the processes to enable the City B's name server as master need to be defined and executed.

There are two major steps to activate the City B's DNS server as the master. First, we can change the zone type from secondary to master in City B name server. Second, we can notify the primary name server of service provider to modify and redirect the source of FIA.CA zone to the City B's name server. The detail execution plan is captured in Section 4.2.2 of this document. Figure 3-18 illustrates the network structure once the change is completed.

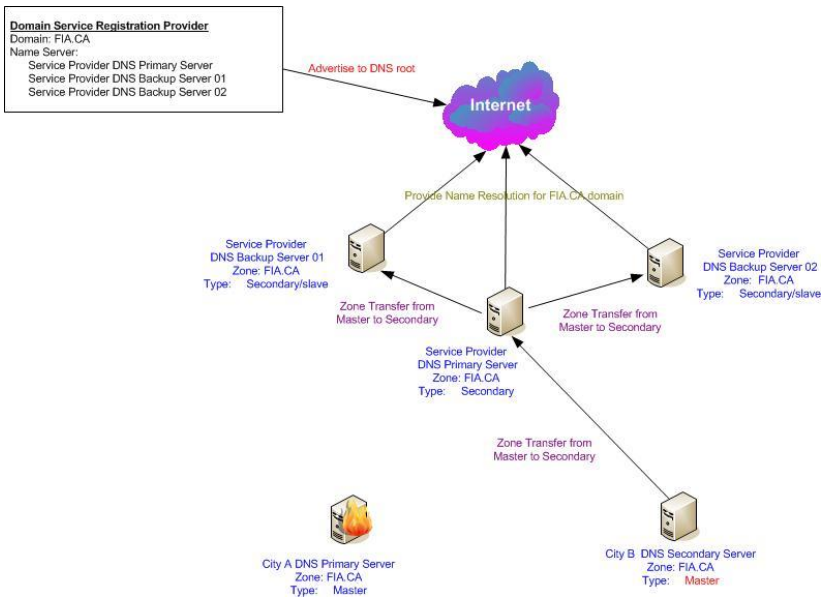

Figure 3-18: FIA External DNS Fail Over

3.5.3 NTP DRP

Corporation wide time synchronization is critical for business success. An increasing number of services are now dependent on accurate time. FIA Active Directory plays the role of corporate authentication and authorization. Microsoft Windows 7 devices authenticate and authorize users through Kerberos protocol against FIA Active Directory. In Kerberos protocol, time difference between client and AD server should not be over 5 minutes. If the difference is over 5 minutes, it often causes unexpected issues.

Other than services needing time synchronization, accurate time is also beneficial to troubleshooting tasks. When there is network incident, checking the log file is always the best starting point. If multiple devices are involved in an incident, all of the devices' log

files need to be verified. If the time is not synchronized between the devices, it increases the difficulty of discovering the cause of the incident from the logs.

FIA built one corporate wide hierarchical network time services infrastructure. In this NTP (Network Time Protocol) service infrastructure, three Cisco network devices are designated as Stratum 3 time sources. Two of the devices are in City A's datacenter. One is in City B's Data Center. All three devices synchronize time with three Internet Stratum 2 time source. The relationship between these three network devices is equal as peer. These three devices provide network time services for the whole corporation. To minimize the traffic through the WAN network, the rest of the network devices are also classified to different Stratums such as 4, 5 and 6.

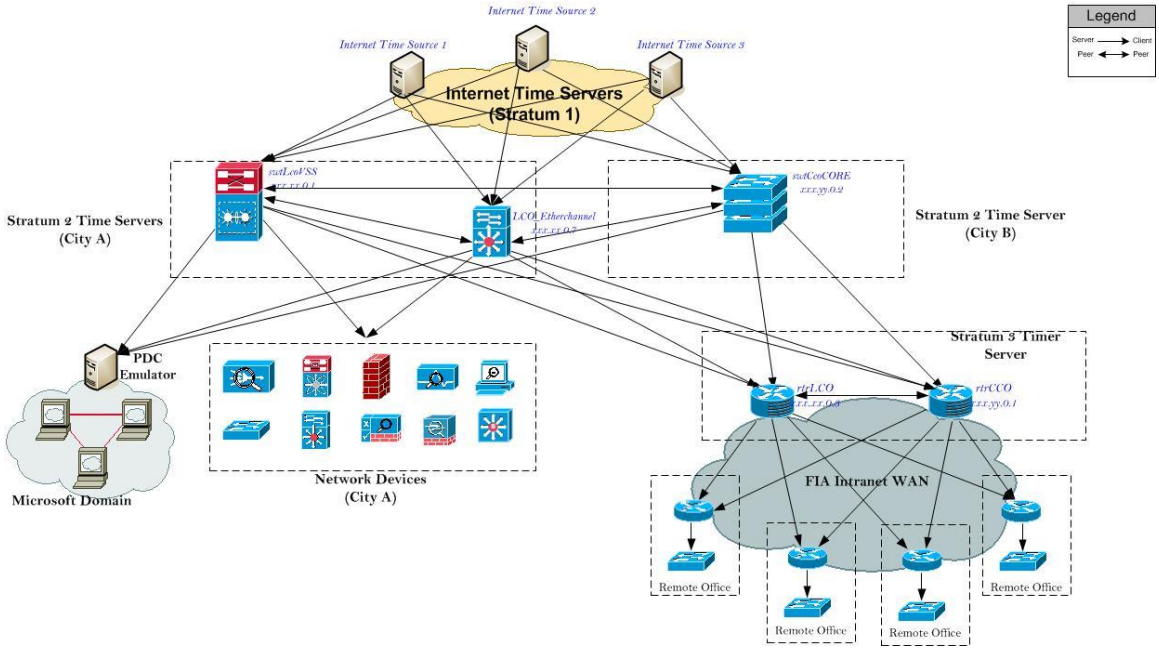

Figure 3-19: FIA NTP Infrastructure

Figure 3-19 shows the infrastructure of the FIA NTP system. For Microsoft Windows domains, four Active Directory Domain Controllers work as NTP server. They all obtain the time from the three Cisco network devices. At the same time, they all provide time service for all Windows devices including server, desktop and laptop. Once the Windows computer is added into the domain, it will automatically synchronize with the four domain controllers. As there are two domain controllers in City A and two in City B, the Windows devices still can get the NTP service through the two domain controllers in City B when City A's office is not available.

The rest of network devices can be partitioned into three parts: network devices in City A's central office, network devices in City B's office and network device in branch offices. Network devices in City A and City B office are all synchronizing with three Stratum 3 time sources. In the other branch offices, only the edge routers synchronize with City A and City B edge routers. The other network devices in the branch offices synchronize with local edge router only. This reduces the traffic through WAN network.

In FIA's internal network, there are three security zones: portal zone, City B DMZ zone, and City A DMZ zone. There are firewalls between the three zones and internal network.

To minimize the security risk, each zone has one network device as NTP server to provide time services for the zone as shown in Figure 3-20:

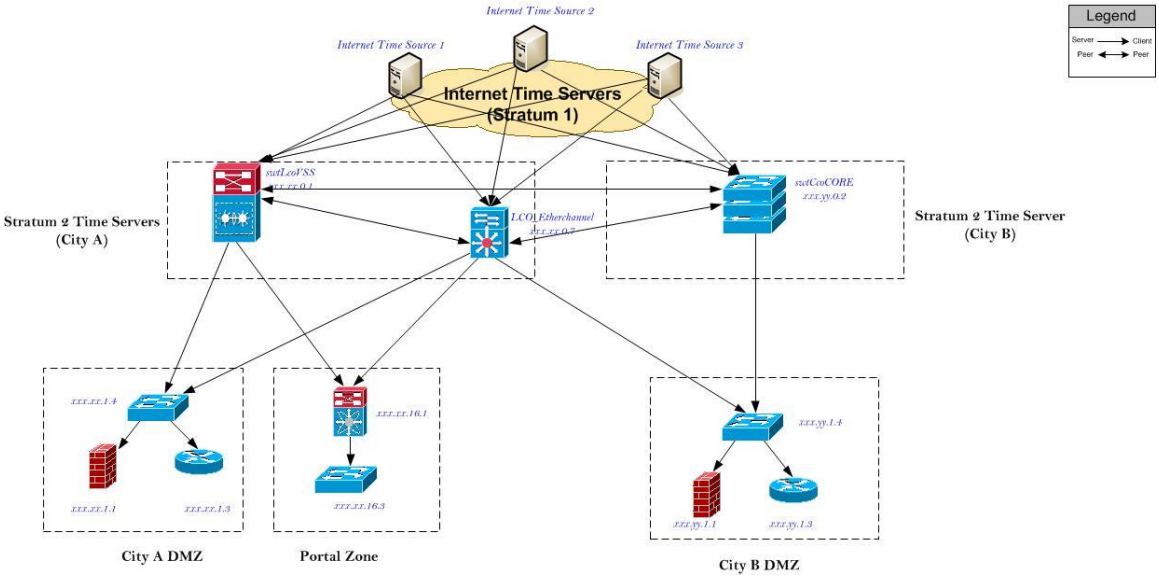

Figure 3-20: FIA NTP Infrastructure of Specific Zones

4. Implementation

FIA DRP Phase I project started in September 2011. The project was divided into several stages: Business Risk Analysis, Requirement Analysis and Design, Implementation and Validation.

The mile stone of the project is as the below table:

No.	Project Stage	Start Time	End Time
1	Business Risk Analysis	9/1/2011	10/15/2011
2	Requirement Analysis and Design	10/16/2011	11/30/2011
3	Implementation	12/1/2011	2/15/2012
4	Validation	2/16/2012	3/23/2012

4.1 Pre-build environment

In the Implementation stage, the initial configuration for the new devices was a challenging task. All the technical resources are located in City A office. In order to ensure time and cost efficiency, it was necessary to build a pre-built environment. In this environment, all the devices were configured the same as the proposed DRP environment in City B. Some basic tests were carried out in the environment to validate the design.

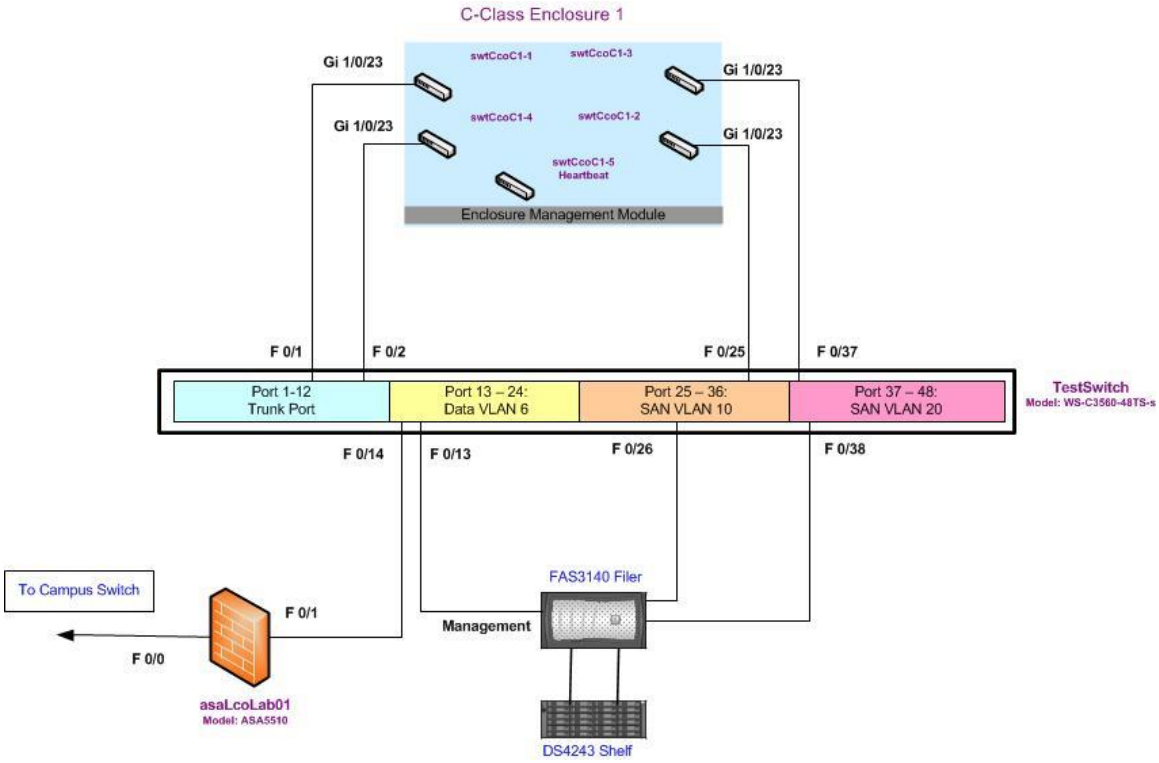

Figure 5-1: Pre-build Environment Topology

The above diagram shows the network topology of the pre-build environment for FIA DRP phase I. One single Cisco Catalyst 4948 switch acts as the core switch. It creates multiple VLANs to simulate City B's network. VLAN 6 is used for the DRP servers' data

connection. VLAN 10 and 20 are used for the servers' storage network connection. VLAN 901 is used for the VMware heartbeat connection. For the Internet connection, one Cisco ASA firewall was deployed within the pre-build environment. The inside interface of the firewall connects to the Cisco Catalyst 4948 switch. The outside interface connects to City A's datacenter network. In the firewall, only the http and https traffic are allowed to pass through. The IP address scope is also the same as the City B datacenter. This minimized the work and effort to change the server's IP addresses after moving to City B's data center.

4.2 Execution Plan

In the DRP project, an execution plan is used to provide step by step instructions to guide the technician to recover the business service. The execution plan reflects the design considerations addressed in the previous section. It is an operational guide and includes all the steps needed to restore services. A good execution plan should be clear and simple. A technician with limited knowledge of the technology used should be able to restore the business following the execution plan. The execution plan also needs to demonstrate the steps to failback the service to primary site when disaster is resolved.

Based on the design, there are two network related execution plans. One is for the internal users to access the Internet. The other one is to active the external DNS server within Service Provider DNS servers.

4.2.1 Internet Access Execution Plan

SYSTEM **Internet Access Services for Internal Users**

OVERVIEW
There are two Internet links for FIA internal users. One Internet link is located in City A's office. And backup Internet link is in City B office. All the Internet traffic is going through City A's Internet link. When disaster happens, all the traffic should go through City B's Internet link instead.

APPLICATIONS
Currently, FIA's network traffic to the Internet is using the default route and is forwarded to xxx.xx.0.4. At same time, there is a backup Internet connection in City B. When City A experience disaster, the default route should point to xxx.yy.0.4 to route all FIA internet traffic through City B.

KEY CONTACTS

Hardware Vendor	Checkpoint Tech Support: +1 (888) 361 5030
	Cisco Tech Support: +1 (800) 553 2447
System Owners	FIA IT Network Team

Scenario 1 Total Loss of City A Data Center
DISASTER RECOVERY PROCEDURE

When Disaster happens, please do the following changes to re-route Internet traffic to City B Internet link:

1. Log on to the City B WAN router (xxx.yy.0.1). Execute following commands in to the configuration:
   #conf t
   #ip route 0.0.0.0 0.0.0.0 xxx.yy.0.4
   # eigrp 2
   # redistribute static

default-metric 10000 1 255 1 1500
#exit
#write mem

Scenario 2 City A Data Center restored
DISASTER RECOVERY PROCEDURE

After City A's datacenter is restored, please do the following changes to re-route Internet traffic to City A Internet link:

1. Log on to the City B WAN router (xxx.yy.0.1). Execute following commands in to the configuration:
   #conf t
   #no ip route 0.0.0.0 0.0.0.0 xxx.yy.0.4
   # eigrp 2
   # no redistribute static
   # no default-metric 10000 1 255 1 1500
   #exit
   #write mem
2. Log on to City A's WAN router (xxx.xx.0.3). Make sure the following command is configured:
   #conf t
   #ip route 0.0.0.0 0.0.0.0 xxx.xx.0.4
   # eigrp 2
   # redistribute static
   # default-metric 10000 1 255 1 1500
   #exit
   #write mem

4.2.2 External DNS Execution Plan

SYSTEM **FIA External DNS Services**
OVERVIEW
FIA external DNS service provides domain name resolution for the FIA.CA domain. External users from Internet access FIA web site by domain lookup through FIA external DNS services

APPLICATIONS
FIA's external DNS service includes 2 DNS servers. One is located in City A's DMZ which hosts the FIA.CA zone as master. The other is located in City B's DMZ which hosts the FIA.CA zone as secondary. The changes on FIA.CA domain can only be made on City A's external DNS server. Through zone transfer, all the records of FIA.CA are copied to Service Provider DNS server as well as FIA City A's external DNS server. External users are connected to one of the Service Provider DNS servers to obtain name resolution service for FIA.CA.
When disaster happens in City A, FIA City B external DNS need to be changed to take master role. At the same time, Service Provider primary DNS server needs to update the setting to allow the zone transfer from the City B external DNS server.

KEY CONTACTS

Hardware Vendor	Checkpoint Tech Support: +1 (888) 361 5030 Cisco Tech Support: +1 (800) 553 2447
System Owners	FIA IT Network Team

Scenario 1 Total Loss of City A Datacenter
DISASTER RECOVERY PROCEDURE

When disaster happens, make the following changes to activate the City B's external DNS server:

1. Log on to the City B's external DNS server. Follow the below steps to make changes
   -- Run DNS management tool
   -- Left click the forward zone FIA.ca
   -- Click Properties
   -- In the General Tap, change the **secondary** to **master**
2. Service Provider DNS Admin login into Service Provider primary DNS server, update FIA.CA zone master settings in the file /etc/named.conf:

```
zone "FIA.ca" {
 type slave;
 file "slave/FIA.ca.zone";
 masters {
       xxx.xx.yy.13;
 };
```

Scenario 2 City A Datacenter restored
DISASTER RECOVERY PROCEDURE

After City A is restored, make the following changes to activate City A's external DNS server:

1. Log on to the City B external DNS server. Follow the below steps to make changes
   -- Run DNS management tool
   -- Left click the forward zone FIA.ca
   -- Click Properties
   -- In the General Tap, change the **master** to **secondary.**
   -- In the General Tap, specify City A's external DNS server as the master

2. Service Provider DNS Admin login into Service Provider primary DNS server and update FIA.CA zone master settings in the file /etc/named.conf:

```
zone "FIA.ca" {
 type slave;
 file "slave/FIA.ca.zone";
 masters {
       xxx.xx.xx.13;
 };
```

5. Disaster Recovery Test (Drill Test)

Practicing in the production environment is an effective way to validate the DRP. After the implementation of FIA DRP Phase I was completed, FIA's IT operation team scheduled a drill test on Saturday March 17th.

In this drill test, the network connections to branch offices and Internet from City A office were taken down to simulate the disaster in the City A office. In this situation, all of the services to the Internet and branch offices were made unavailable. FIA Disaster Recovery Team followed the DRP execution plan and process to restore the services in City B. Once the DRP was completed, the FIA business test team validated the recovered systems. During this test, three critical business services were recovered: web server, finance system and Exchange system of FIA. The key support systems were also recovered during the drill test including network, Internet, DHCP, DNS, AD and Tape backup.

Once all the validation work was completed, the network connection of City A office was re-activated. The FIA Recovery Team then failed back to City A office. With all the systems validated by the business test team, the drill test was finished.

The whole test took about 8 hours. It verified the correction of the DRP plan. It also added valuable experience to the team members of the DRP plan.

The results of this drill test showed that the DRP plan of FIA DRP phase I is effective and practical. It addresses the entire current FIA DRP requirement.

Based upon the results from the drill testing, FIA IT Operation Manager announced the project as being successful.

6. Improvement for Next Phase

DRP plan is an ongoing process. Technology is constantly evolving over time, and the DRP plan also needs to be updated with new solutions. Although FIA DRP phase I meets the current business requirement, it is not the end point. There is still room for improvement within the network infrastructure.

Data Center Infrastructure Upgrade

FIA's current datacenter infrastructure in both City A and City B is not upgraded to the industry's recommended standard. Many devices in the datacenter are at the end of their support period from the manufacturers. However, they are still in use. Not only does it increase the business risk, but it also restricts the potential to improve the RPO and RPT of the DRP plan.

FIA's datacenter network infrastructure at City B has less capacity and backbone bandwidth compared to the datacenter at City A. In City A's datacenter, the core switch provides 10 Gbps ports for distribute and access switch. Only 1 Gbps port is available in City B's datacenter core switch. It is necessary to assess the traffic volume before adding services onto the DRP site. It can help avoid potential performance issues when new services are restored during the disaster period.

Optimization of WAN Link

In FIA DRP phase I, only three business services are covered. The current bandwidth of the WAN link can accommodate the traffic generated by both regular and replication access. However, this is not the case once more critical services are added into the DRP plan. It is foreseeable that more varied traffic will be generated between two datacenters through WAN link. To achieve higher RTO and RPO, the current WAN link needs to be reviewed. Optimizing current WAN link and installing a dedicated replication WAN link are two of the available solutions. Choosing one or both solutions will depend on traffic characteristics of the services.

Routing Technology Improvement

Cisco EIGRP is a mature and dynamic routing protocol. Its advantages over other products include simple configuration, fast convergence, less hardware requirement and easy management. However it does not meet the new DRP solution requirement in full. Virtual datacenter is an emerging DRP solution. The virtual datacenter is built upon multiple physical datacenters. In a specific time instant, the virtual datacenter may be located in one physical datacenter or multiple physical datacenters. When one physical data center fails, the virtual data center can be easily moved to another physical datacenter with low down time. To achieve this virtual datacenter, Cisco has developed a new protocol named LISP (Locater/ID Separation protocol) to relocate the virtual data center IP address from one physical datacenter to another datacenter automatically. For this purpose, FIA intranet routing architecture needs to be improved in the next DRP phase.

IPv6 Consideration

IPv6 is gaining popularity over time. FIA doesn't have the urgent need to deploy IPv6 in its internal network. However the number of services that require IPv6 support is increasing. Microsoft's DirectAccess VPN (Virtual Private Network) solution is one of the many services that incorporate IPv6 as standard. IPv6 design should be included in the next DRP phase.

Appendix A: Acronym Glossary

AD	Microsoft Active Directory
AS	Autonomous System
BCP	Business Continues Plan
CAS	Client Access Server
CPE	Customer Premises Equipment
DAG	database availability group
DHCP	Dynamic Host Control Protocol
DMZ	Demilitarized Zone
DNS	Domain Name System
DRP	Disaster Recovery Plan
DSRP	Domain Service Registration Provider
DUAL	Diffusing Update Algorithm
EIGRP	Enhanced Interior Gateway Routing Protocol
FIA	Financial Institution Alfa
Gbps	Gigabits Per Second
GUI	Graphical User Interface
HTTPS	Hypertext Transfer Protocol Secure
ICMP	Internet Control Message Protocol
iLO	Integrated Lights-Out
IP	Internet Protocol
iSCSI	Internet Small Computer System Interface
LISP	Locater/ID Separation protocol
Mbps	Megabits Per Second
MDT	Microsoft Deployment Tool
msec	Millisecond
NICs	Network Interface Cards
NTP	Network Time Protocol
PDC	Primary Data Center
RPO	Recovery Point Object
RR	Resource Records
RTO	Recovery Time Object
SAN	Storage Area Network
SCCM	System Center Configuration Manager
SCSI	Small Computer System Interface
SDC	Secondary Data Center
SMTP	Simple Mail Transfer Protocol
SLA	Service Level Agreement
SOA	Start of Authority
SPAN	Switched Port Analyzer
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
VLAN	Virtual Local Area Network
VM	Virtual Machine
VPN	Virtual Private Network
WAN	Wide Area Network