

Microsoft Exchange Server 2007—Cisco, EMC, and VMware Multisite Data Center Design

January 10, 2009

About the Document

This document provides design and configuration guidance for site and server load balancing, Secure Sockets Layer (SSL)- offload and WAN optimization, SAN and HA/DR solutions in a virtualized Microsoft Exchange Server 2007 environment. An overview of the various Microsoft Exchange Server 2007 roles and operations are discussed to provide the reader some context as to how the application environment is impacted in a multisite data center design.

Audience

This document is intended for network, SAN and Exchange administrators, and architects who need to understand both the basics of a Microsoft Exchange environment and the design and configuration options for providing advanced network services for Microsoft Exchange Server 2007.

Document Objectives

The objective of this document is to provide customers guidance on how to use a jointly deployed Cisco, EMC, and VMware multisite data center design to support a Microsoft Exchange Server 2007 environment. The document is not meant to introduce the reader to basic Cisco, EMC, or VMware configurations; nor is it meant to be a resource to learn the details of Microsoft Exchange Server 2007. The reader must be familiar with the basic Cisco, EMC, and VMware technology concepts and products as well as the basics of Microsoft Exchange Server 2007 components, roles, and deployment scenarios as documented by Microsoft Corporation. The prerequisite knowledge can be acquired through many documents and training opportunities available both through Cisco and Microsoft. For a list of recommended information resources list that readers may find useful in these areas of interest, refer to [Reference Documents](#), page 92.



Corporate Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

Copyright © 2007 Cisco Systems, Inc. All rights reserved.

Solution Overview

Microsoft Exchange, considered the backbone of enterprise messaging environments, is growing at a rapid pace. As the number of users and departments grow, so does the requirement for a scalable, reliable, and cost-effective Microsoft Exchange architecture to meet the new business requirements. While enterprises prepare to meet demands for greater collaboration, quicker access to applications, and compliance with ever-stricter regulatory compliance, they are hampered by problems relating to power and cooling, efficient asset usage, escalating security and provisioning needs, and business continuance. Customers understand that temporary growth is no longer viable and appreciate the benefits they can gain from data center consolidation and virtualization. A virtualized data center infrastructure enables customers to efficiently service a Microsoft Exchange collaborative environment.

Microsoft Exchange messaging environments are also growing in complexity, and user requirements are increasingly demanding. Additionally, the manner in which Microsoft Exchange is used to support business operations has changed, and it is now even more business-critical than ever before.

Without a comprehensive plan for a Microsoft Exchange 2007 deployment, a company could face the following risks:

- Loss of revenue
- Missed business opportunities
- Compliance-related fines
- Loss of data

With these potential risks of a poorly implemented solution, our market analysis revealed the following critical business factors that affect Microsoft Exchange 2007 installations:

- Capital cost reduction through consolidation
- Operating expense reduction through streamlining operations management
- Risk reduction through validated compliance
- Risk reduction through security (including spam and virus filtering needs)

The design presented in this document enables customers considering a Microsoft Exchange 2007 deployment to capitalize on the benefits of a virtualized platform which offers consolidation, simplified management, compliance, and security. The solution is based on VMware ESX 3.5, EMC storage and replication capabilities that provide disaster-recovery protection, and advanced server, storage connectivity and application optimization through Cisco technology.

Solution Components and Topology

Figure 1 depicts the Microsoft Exchange Server 2007 solution topology tested, where two distinct data centers (Data Center 1 and Data Center 2) are deployed using Cisco's infrastructure design best practices. Note that each site provides local redundancy, scalability, and security for the applications it hosts. A multisite solution should simply extend the functionality of a single-site and should not compromise the integrity of either.

At each site in Figure 1, the following products and technologies are used:

EMC:

- The EMC Symmetrix DMX™ is a high-end storage solution used for storing all of the VMware ESX Virtual Machine data as well as the Microsoft Exchange 2007 Mailbox database and log data.

- EMC RecoverPoint provides an appliance-based solution for continuous data protection and continuous remote replication for on-demand protection and recovery of the Microsoft Exchange 2007 Mailbox role. EMC RecoverPoint ensures continuous replication to the second data center location without impacting performance and is tightly integrated with VMware Site Recovery Manager (SRM). RecoverPoint works in conjunction with the Cisco SANtap with the Storage Services Module (SSM) in the MDS to provide seamless storage replication for the Exchange 2007 Mailbox VM and associated database and log data. Along with the Cisco SANtap/SSM/MDS solution, EMC RecoverPoint replicates from FibreChannel (FC) at the block level to a remote recovery array over the data center-to-data center WAN IP network.

Cisco:

- The SAN is connected together via the Cisco MDS Fabric Switch (green lines in Figure 1). The EMC RecoverPoint and DMX storage arrays are connected to the Cisco MDS as are the VMware ESX hosts and Nexus 5000 for Unified Fabric connectivity.
- The data center access layer provides the network entry point for hosts. A range of Cisco products are used to provide access for Gigabit attached hosts such as the Catalyst 4948, 4900M, 6500 and Nexus 7000. The Cisco Nexus 5000 is used for Unified Fabric in the access layer. The VMware ESX hosts that have Converged Network Adapters (CNA) installed connect to the Nexus 5000 for both Ethernet and Storage access (red lines in Figure 1).
- The data center aggregation layer has Cisco Catalyst 6500s with Sup720s and is the aggregation point for all access layer connections. In addition to being the Layers 2 and 3 boundary point, the aggregation layer contains a variety of service modules to include the Cisco Application Control Engine (ACE) and the Firewall Services Module (FWSM). In this document the Cisco ACE provides the SLB and SSL-offload services to the Exchange 2007 Client Access Server (CAS) role.
- The data center core layer has Cisco Nexus 7000s. The core is the connecting point between other places in the network (PIN) such as the Internet edge, WAN, and campus.
- The Internet edge has the Cisco Global Site Selector (GSS) deployed to provide intelligent site selection for both the Exchange 2007 CAS role and also the IronPort C-Series for SMTP relay. The ASA is used to provide perimeter firewall security services for the Exchange 2007 CAS and the IronPort C-Series. Finally, the Ironport C-Series secure email appliance is used to provide Exchange 2007 messaging security, conformance and relaying.
- The branch site is used to validate Outlook 2007 Messaging Application Programming Interface (MAPI) and Outlook Web Access (OWA) connections. Both the MAPI and OWA connections are optimized using the Cisco Wide Area Application Services (WAAS). There are Cisco Wide Area Application Engines (WAE) at the branch. WAN edge and the Central Manager (CM) is located in the data center access layer.
- Cisco WAAS is deployed in the data center using both Web Cache Communication Protocol (WCCP) and Inline traffic interception methods. WCCP is configured on the WAN router that supports branch connectivity to optimize application traffic to and from the branch. Inline interception is used between the data center core switches and the WAN router that supports the DC-to-DC WAN link. This is used to support the storage replication traffic optimization between data centers.

VMware:

- Every Exchange 2007 server role in this solution is virtualized using VMware ESX 3.5 Update 3 and Virtual Center 2.5. A VMware ESX cluster contains multiple ESX hosts and the High Availability (HA) and Distributed Resource Scheduler features of Virtual Center are used to manage the availability and performance of each Exchange 2007 VM.

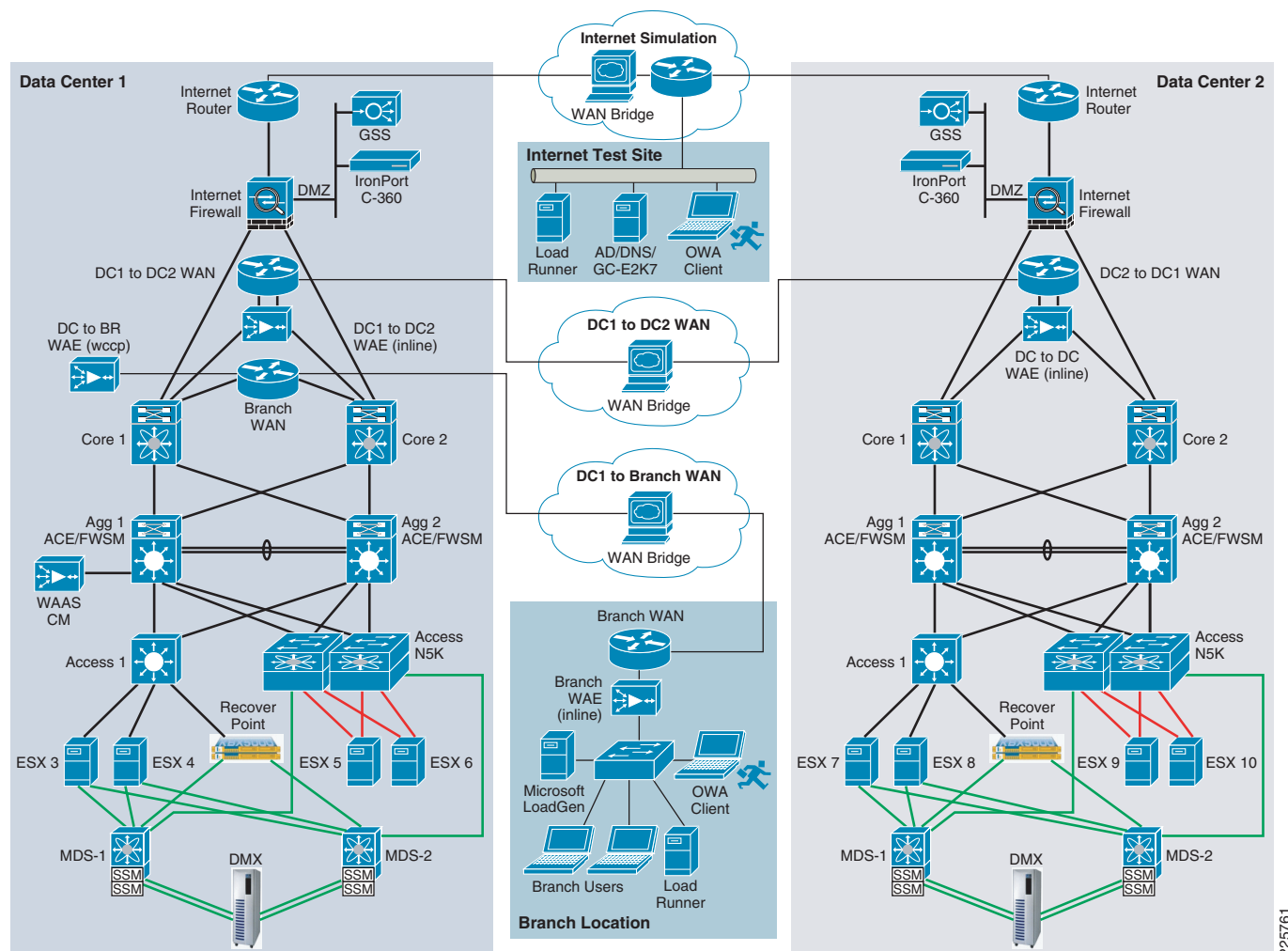
- The VMware Site Recovery Manager (SRM) 1.0 product is used along with EMC RecoverPoint to ensure that the Exchange 2007 Mailbox role and its associated SAN-attached LUNs are replicated to Data Center 2. Once a decision is made to activate the SRM recovery plan for the MBX role, the exact replica (including database and log information) is activated at Data Center 2 and MBX access is restored.

Microsoft:

- Microsoft Windows Server 2008 Data Center Edition is used as the server operating system on each VM, unless otherwise noted.
- Microsoft Windows Vista Enterprise is used as the client operating system.
- Microsoft Exchange Server 2007 SP1:
 - Hub Transport Role— Performs the central role for all intelligent message routing, delivery and control within and outside of the organization. Multiple Hub Transports are deployed at both sites to offer redundancy and HT load-balancing services. The HT roles have connections to the IronPort C-Series for SMTP relay purposes.
 - CAS Role— Provides messaging access to a variety of client endpoints to include OWA, Outlook Anywhere, and ActiveSync clients. Multiple CAS are deployed at both sites to offer redundancy. The Cisco ACE is used to provide Server Load Balancing and SSL-offload of the CAS servers.
 - Mailbox Role— The Mailbox (MBX) role is the database for all user messaging data; it provides access to MAPI-based clients such as Outlook. The MBX is deployed at Data Center 1 and is considered the primary location for mail services. The Mailbox server and associated storage is replicated (see below for technologies used for this replication) to the Data Center 2 site for DR purposes.

This document discusses each of the areas defined in [Figure 1](#) to provide a better understanding of the application and the network deployed to support it.

Figure 1 Solution Topology



EMC Technology Overview

DMX 1000

- The DMX 1000 provides Enterprise class storage with up to 64 direct nonblocking data paths in the DMX1000 and up to 35.2Gb/s aggregate of internal bandwidth. The Symmetrix DMX1000 system can support up to four slots in the midplane dedicated to global memory directors and a maximum of 128 Gigabits of global memory. A minimum of two and a maximum of four global memory director boards are required for the DMX1000 system configuration.
- A 2Gbps FibreChannel Drive Infrastructure provides multiple scalable channel directors, disk directors, and global memory directors. A maximum number of 144 disk drives and redundant loop are provided for a extremely redundant and scalable design.
- Testing included LUNs provided by the DMXs to redundantly connected ESX hosts.

RecoverPoint

EMC RecoverPoint is an appliance-based solution that copies write I/Os from the Exchange 2007 MBX server and transport that data to second data center synchronously or asynchronously. This solution allows the second data center to provide a different tiered storage or even different storage from another vendor. The data is captured from a SAN infrastructure and sent to the appliance. Then the RecoverPoint appliance sends the data to another RecoverPoint appliance at the secondary data center through the data DC-to-DC interconnect, which is an IP-based infrastructure. The SAN switches do not need to have IP ports since the RecoverPoint appliance provides the FibreChannel (FC) to IP encapsulation/decapsulation. RecoverPoint works over an existing WAN without requiring expensive edge devices with only a 3Mbps minimum bandwidth requirement for replication across an IP WAN. RecoverPoint appliances do not require identical storage equipment at a recovery site to do native array replication such as with SRDF.

EMC RecoverPoint provides various value added propositions for Continuous Data Replication (CDP) and Continuous Remote Replication (CRR) within the solution. This document focuses only on the integration of RecoverPoint with VMware Site Recovery Manager (SRM) to protect the Exchange 2007 Mailbox role. SRM currently only supports Continuous Remote Replication and not CDP. RecoverPoint Appliances provides redundancy by clustering appliances. Two to eight RecoverPoint Appliance (RPAs) are supported per site. With a maximum replicated LUN size of 2TB, a maximum of 30 constancy groups per RPA, and a maximum of 2048 replicated LUNS in single cluster, RecoverPoint offers the SAN administrator a highly powerful set of tools to handle many different replication requirements including integration with VMware Site Recovery Manager.

Cisco SANTap technology to support RecoverPoint is provided through the Cisco MDS Storage Services Module (SSM), which has 32 FC switching ports and 8 embedded ASICs that provide SANTap capabilities. SANTap currently supports proxy-mode, which requires the application server to reside in a different vSAN as the storage. For Exchange 2007 Mailbox servers or other Exchange 2007 roles deploying SANTap, there is a one-time disruption to the application. To mitigate this disruption, servers with dual paths to the storage can implement SANTap on one path at a time, to minimize application disruption. SANTap scalability with the Switch Service Module (SSM) and Cisco MDS are well supported to match growing storage replication needs for the Exchange 2007 environment.



Note

For best practice and design guides, refer to the EMC's website at <http://powerlink.emc.com> and the *Reference EMC® RecoverPoint Deploying RecoverPoint with SANTap Technical Notes P/N 300-004-387 Rev A07*.

Cisco Technology Overview

This section provides an overview of the main Cisco products and technologies used in this design. The following products are addressed:

- Cisco Application Control Engine (ACE) provides intelligent Server Load Balancing (SLB) for Microsoft Exchange 2007 Server components. Cisco ACE optimizes server resources by ensuring these resources are load balanced intelligently with advanced algorithms and user-tunable parameters. These services increase application availability and optimize server resources. The Cisco ACE also provides SSL and TCP offload capabilities for the CAS role.
- With multiple data centers, the Cisco ACE Global Site Selectors (GSSs) are used in conjunction with Cisco ACE SLB to provide workload distribution, disaster recovery, failover protection, and Domain Name System (DNS) offloading. Optimal site load balancing is achieved by having the Cisco ACEs within the data center in active communications with the GSSs (outside the data center) to ensure optimal access across data centers.

- Cisco Wide Area Application Services (WAAS) is a comprehensive WAN optimization solution that accelerates MS Exchange MAPI and OWA connections over the WAN, delivers video to the branch office, and provides local hosting of branch-office IT services. Cisco WAAS allows IT departments to centralize applications and storage in the data center while maintaining LAN-like application performance.
- Cisco Adaptive Security Appliance (ASA) provides application-aware advanced firewall and VPN security services at the perimeter of the network

Cisco IronPort C Series appliances reside at the edge of the network behind an ASA firewall. These devices protect Microsoft Exchange servers from virus attacks, spam, and other malware before the email message is injected into the corporate network by ensuring that unwanted email messages are automatically filtered without user intervention.



Note

The generic design and configuration details for the Cisco data center access, aggregation and core layers are not specific to Microsoft Exchange 2007 and therefore are not included in this document. Refer to the Cisco Validated Designs (CVD) for more information on the recommended Cisco infrastructure deployment options: <http://www.cisco.com/go/designzone>.

ACE

The Cisco ACE provides a highly available and scalable data center solution from which the Microsoft Exchange Server 2007 application environment can benefit. Currently, the Cisco ACE is available as an appliance or integrated service module in the Cisco Catalyst 6500 platform. The Cisco ACE features and benefits include the following:

- Device partitioning (up to 250 virtual ACE contexts)
- Load balancing services (up to 16 Gbps of throughput capacity and 345,000 Layer 4 connections/second)
- Security services through deep packet inspection, access control lists (ACLs), unicast reverse path forwarding (uRPF), Network Address Translation (NAT)/Port Address Translation (PAT) with fix-ups, syslog, and so on
- Centralized role-based management via Application Network Manager (ANM) GUI or CLI
- SSL-offload (up to 15,000 SSL sessions through licensing)
- Support for redundant configurations (intra-chassis, inter-chassis, and inter-context)

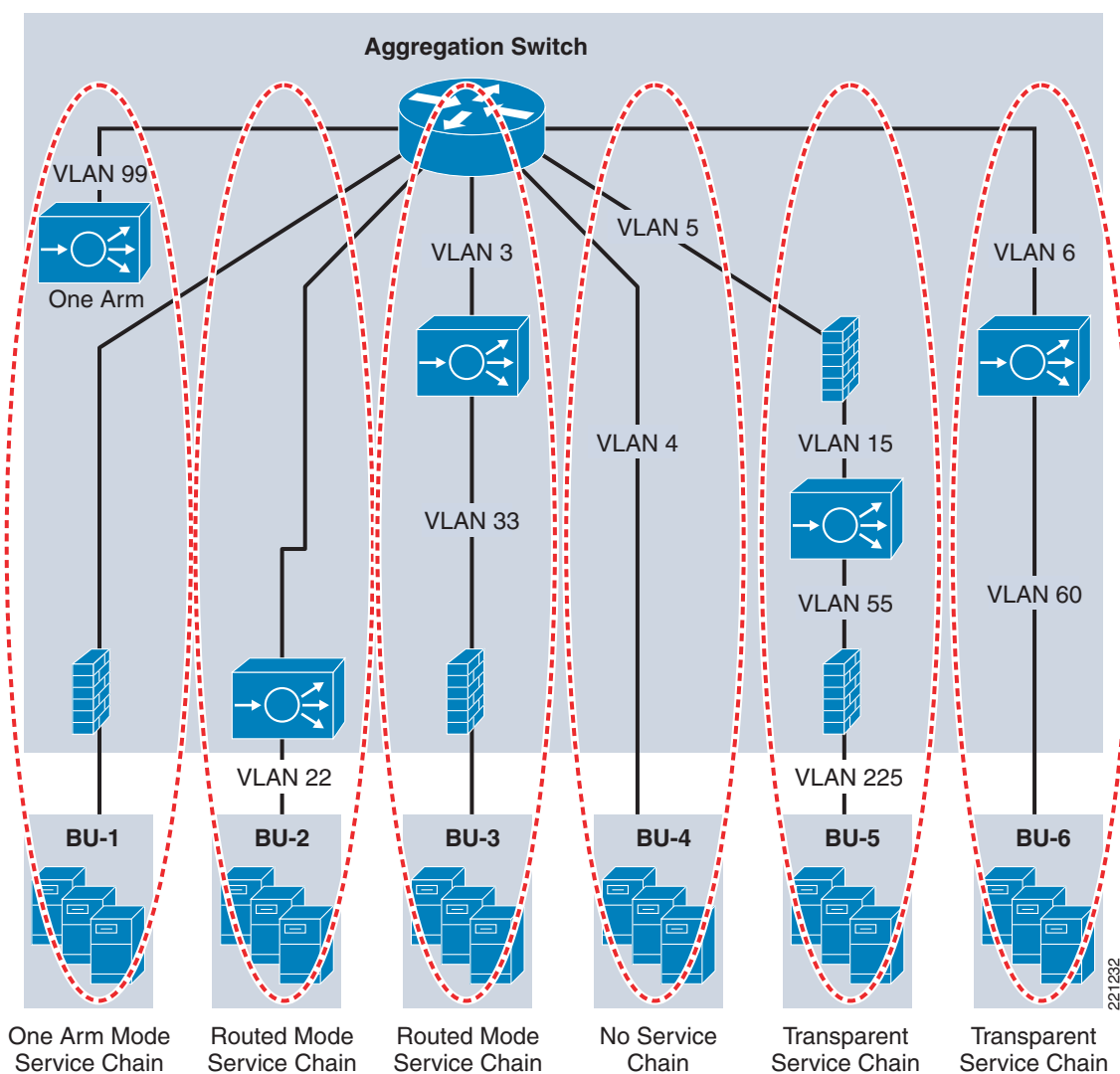
ACE Virtualization

Virtualization is a prevalent trend in the enterprise today. From virtual application containers to virtual machines, the ability to optimize the use of physical resources and provide logical isolation is gaining momentum. The advancement of virtualization technologies includes the enterprise network and the intelligent services it offers.

The Cisco ACE supports device partitioning where a single physical device may provide multiple logical devices. This virtualization functionality allows system administrators to assign a single virtual ACE device to a business unit or application to achieve application performance goals or service-level agreements (SLAs). The flexibility of virtualization allows the system administrator to deploy network-based services according to the individual business requirements of the customer and technical requirements of the application. Service isolation is achieved without purchasing another dedicated appliance that consumes more space and power in the data center.

Figure 2 shows the use of virtualized network services afforded through the Cisco ACE and Cisco Firewall Services Module (FWSM). In Figure 2, a Cisco Catalyst 6500 housing a single Cisco ACE and FWSM supports the business processes of five independent business units. The system administrator determines the application requirements and assigns the appropriate network services as virtual contexts. Each context contains its own set of policies, interfaces, resources, and administrators. The Cisco ACE and FWSMs allow routed, one-arm, and transparent contexts to coexist on a single physical platform.

Figure 2 Service Chaining via Virtualized Network Services



Note

For more information on ACE virtualization, see the *Application Control Engine Module Documentation* at the following URL:

http://www.cisco.com/en/US/products/ps6906/tsd_products_support_model_home.html

SSL-Offload

The Cisco ACE is capable of providing secure transport services to applications residing in the data center. The Cisco ACE implements its own SSL stack and does not rely on any version of Open SSL. The Cisco ACE supports TLS 1.0, SSLv3, and SSLv2/3 hybrid protocols. There are three SSL relevant deployment models available to each ACE virtual context:

- *SSL termination*—Allows for the secure transport of data between the client and ACE virtual context. The Cisco ACE operates as an SSL proxy, negotiating and terminating secure connections with a client and a non-secure or clear text connection to an application server in the data center. The advantage of this design is the offload of application server resources from taxing the CPU and memory demands of SSL processing, while continuing to provide intelligent load balancing.
- *SSL initiation*—Provides secure transport between the Cisco ACE and the application server. The client initiates an unsecure HTTP connection with the ACE virtual context, the Cisco ACE acting as a client proxy negotiates an SSL session to an SSL server.
- *SSL end-to-end*—Provides a secure transport path for all communications between a client and the SSL application server residing in the data center. The Cisco ACE uses SSL termination and SSL initiation techniques to support the encryption of data between client and server. Two completely separate SSL sessions are negotiated, one between the ACE context and the client, the other between the ACE context and the application server. In addition to the intelligent load balancing services the Cisco ACE provides in an end-to-end SSL model, the system administrator may choose to alter the intensity of data encryption to reduce the load on either the frontend client connection or backend application server connection to reduce the SSL resource requirements on either entity.

SSL URL Rewrite Offload

The Cisco ACE is capable of inserting or deleting HTTP header information for connections it is sustaining. This capability is highly useful when an application server responds with a HTTP 302 or “Moved Temporarily” response to a client's HTTP GET or HEAD request. The HTTP 302 response usually indicates a new HTTP LOCATION URL for the client to access. Modifying the HTTP LOCATION value for a secure connection is known as SSL URL rewrite. The SSL URL Rewrite feature allows the system administrator to alter the HTTP LOCATION value returned to the client resulting in granular control of the application's session flow and persistence in the data center.

SSL Session ID Reuse

SSL session ID reuse allows the client and server to reuse the secret key negotiated during a previous SSL session. This feature generally improves the volume of SSL sessions that an SSL server or SSL proxy can effectively maintain. Clients residing with remote connectivity, for instance across a WAN, generally benefit from this feature. The SSL negotiation load is effectively reduced on the SSL proxy server while simultaneously improving the user experience as key negotiation is a rather lengthy process. The Cisco ACE may maintain the SSL session ID indefinitely or up to 20 hours with a timeout configuration.

It should be noted that SSL ID reuse does not compromise the security of the data center. The ID reuse feature only acknowledges that a secret key already exists between the client and server. Nonetheless, the client must use this key for the application server to receive data from the client. The security resides in the secret key, not the SSL session ID.

Session Persistence

Session persistence is the ability to forward client requests to the same server for the duration of a session. Microsoft supports session persistence for their Microsoft Exchange environment via the following methods:

- Source IP sticky
- Cookie sticky

The Cisco ACE supports each of these methods, but given the presence of proxy services in the enterprise, Cisco recommends using the cookie sticky method to guarantee load distribution across the serverfarm wherever possible as session-based cookies present unique values to use for load balancing. The following trace shows the **sessionid** cookie inserted into the client's Microsoft Exchange request through the **Set-Cookie** command from the server. The cookie named *ACE-Insert* is inserted by ACE based on the configuration settings discussed later in the paper. Using cookie insert, the ACE inserts the cookie on behalf of the server upon the return request, so that the ACE can perform persistent loadbalancing even when the servers are not configured to set cookies.

```
HTTP/1.1 302 Moved Temporarily
Set-Cookie: ACE-Insert=R431233881; path=/
Content-Length: 0
Location: http://owa.esa.com/owa/auth/logon.aspx?url=http://owa.esa.com/owa/&reason=0
Set-Cookie: sessionid=; path=/; expires=Thu, 01-Jan-1970 00:00:00 GMT
Set-Cookie: cadata=; path=/; expires=Thu, 01-Jan-1970 00:00:00 GMT
Date: Tue, 06 Jan 2009 00:42:08 GMT
```

In addition, the Cisco ACE supports the replication of sticky information between devices and their respective virtual contexts. This provides a highly available solution that maintains the integrity of each client's session.

Allowed Server Connections

Enterprise data centers should perform due diligence on all deployed server and network devices, determining the performance capabilities to create a more deterministic, robust, and scalable application environment. The Cisco ACE allows the system administrator to establish the maximum number of active connections values on a per-server basis and/or globally to the serverfarm. This functionality protects the end device, whether it is an application server or network application optimization device such as the WAE.

Route Health Injection

Route Health Injection (RHI) allows the Cisco ACE to advertise host routes associated with any number of virtual IP addresses hosted by the device. The injection of the host route to the remaining network offers Layer 3 availability and convergence capabilities to the application environment.

KAL-AP UDP Agent

The Cisco ACE supports the KeepAlive-Appliance Protocol (KAL-AP) via a local UDP agent. This agent responds to KAL-AP queries from site selectors, such as the Cisco Global Site Selector, to provide the status and workload associated with one or more virtual IP addresses maintained by an ACE virtual context. The KAL-AP agent supports both domain and tagged formed queries. Tagged formed queries allow the verification of VIP state across NAT devices, such as firewalls or routers, and multiple ports for the same virtual IP address. This real-time information provides a more robust and accessible application as load and availability information may be used to distribute traffic intelligently across multiple enterprise sites.

Health Monitoring

The Cisco ACE device is capable of tracking the state of a server and determining its eligibility for processing connections in the serverfarm. The Cisco ACE uses a simple pass/fail verdict but has many recovery and failures configurations, including probe intervals, timeouts, and expected results. Each of these features contributes to an intelligent load-balancing decision by the ACE context.

Following are the predefined probe types currently available on the ACE module:

- ICMP
- TCP
- UDP
- Echo (TCP/UDP)
- Finger
- HTTP
- HTTPS (SSL Probes)
- FTP
- Telnet
- DNS
- SMTP
- IMAP
- POP
- RADIUS
- Scripted (TCL support)

Note that the potential probe possibilities available via scripting make the Cisco ACE an even more flexible and powerful application-aware device. In terms of scalability, the Cisco ACE module can support 1000 open probe sockets simultaneously.

Application Control Engine Global Site Selector

Overview

The Cisco ACE Global Site Selector (GSS) is an appliance that offers failover protection through Global Server Load Balancing (GSLB). The Cisco GSS device allows the enterprise to distribute and balance workload across multiple sites, providing the following benefits:

- Work-load distribution
- Disaster recovery and failover protection
- Improved user experience
- DNS offload

The Cisco GSS becomes part of the enterprise's DNS routing hierarchy as the authoritative DNS server for those services under its domain. The Cisco GSS intelligently resolves DNS requests with the additional knowledge of the site's availability and the associated application's state. This knowledge is gained from tight integration with load-balancers such as the Cisco Content Services Switch (CSS), Cisco Content Switch Module (CSM), and the Cisco ACE. Each of these load-balancers monitor the

state of local application servers and communicate this information to the Cisco GSS, where a global enterprise aware decision can be made. Currently, the Cisco GSS can support approximately 4,000 virtual IP addresses. The Cisco GSS includes the following factors prior to responding to a DNS request:

- Availability
- Proximity
- Load
- Source of the request (DNS proxy)
- Preference

**Note**

The Cisco GSS device may also monitor individual servers, IOS SLB devices, DRP-enabled routers, Cisco's Local Director, and Cisco cache engines.

Keepalives

The Cisco GSS uses keepalives to determine the state of a particular VIP under its domain. The Cisco GSS supports the following keepalive types:

- ICMP
- TCP
- HTTP HEAD
- KeepAlive-Appliance Protocol (KAL-AP)
- Scripted Keepalives
- Name Server

These keepalive types can be used individually or in a multiport group to determine the status of a virtual IP address. As a rule, the Cisco GSS does not respond to a DNS query with a VIP that has been declared inactive.

The KAL-AP keepalive is particularly useful when the Cisco network load-balancing technology is present. The Cisco GSS queries the load-balancer at each site for VIP state and load information. The detailed response received by the Cisco GSS from the network load-balancer can be used to distribute load efficiently across sites.

**Note**

The keepalive timers may be adjusted to establish an acceptable failure window for the enterprise.

Wide Area Application Engine (WAAS)

To appreciate how the Cisco WAAS provides WAN and application optimization benefits to the enterprise, consider the basic types of centralized application messages that are transmitted between remote branches. For simplicity, two basic types are identified:

- *Bulk transfer applications*—Transfer of files and objects, such as FTP, HTTP, and IMAP. In these applications, the number of round-trip messages may be few, and may have large payloads with each packet. Examples include web portal or thin client versions of Oracle, SAP, Microsoft (SharePoint, OWA) applications, e-mail applications (Microsoft Exchange, Lotus Notes), and other popular business applications.

- *Transactional applications*—High number of messages transmitted between endpoints. Chatty applications with many round-trips of application protocol messages that may or may not have small payloads.

The Cisco WAAS uses the technologies described in the following subsections to provide a number of features, including application acceleration, file caching, print service, and DHCP to benefit both types of applications.

Advanced Compression Using DRE and Lempel-Ziv Compression

Data Redundancy Elimination (DRE) is an advanced form of network compression that allows the Cisco WAAS to maintain an application-independent history of previously-seen data from TCP byte streams. Lempel-Ziv (LZ) compression uses a standard compression algorithm for lossless storage. The combination of using DRE and LZ reduces the number of redundant packets that traverse the WAN, thereby conserving WAN bandwidth, improving application transaction performance, and significantly reducing the time for repeated bulk transfers of the same application.

Transport File Optimizations

The Cisco WAAS Transport File Optimizations (TFO) uses a robust TCP proxy to safely optimize TCP at the WAE device by applying TCP-compliant optimizations to shield the clients and servers from poor TCP behavior because of WAN conditions. The Cisco WAAS TFO improves throughput and reliability for clients and servers in WAN environments through increases in the TCP window sizing and scaling enhancements as well as implementing congestion management and recovery techniques to ensure that the maximum throughput is restored if there is packet loss.

Common Internet File System Caching Services

Common Internet file system (CIFS), used by Microsoft applications, is inherently a highly chatty transactional application protocol where it is not uncommon to find several hundred transaction messages traversing the WAN just to open a remote file. The Cisco WAAS provides a CIFS adapter that can inspect and to some extent predict what follow-up CIFS messages are expected. By doing this, the local WAE caches these messages and sends them locally, significantly reducing the number of CIFS messages traversing the WAN.

Print Services

The Cisco WAAS provides native SMB-based Microsoft print services locally on the WAE device. Along with CIFS optimizations, this allows for branch server consolidation at the data center. Having full-featured local print services means less traffic transiting the WAN. Without the Cisco WAAS print services, print jobs are sent from a branch client to the centralized server(s) across the WAN, and then back to the branch printer(s), thus transiting the WAN twice for a single job. The Cisco WAAS eliminates the need for either WAN trip.

Cisco WAAS Replication Accelerator Mode for EMC Recover Point Optimization

Cisco WAAS uses several WAN optimization techniques, including TCP optimization called transport flow optimization (TFO), caching and compound compression known as data redundancy elimination (DRE), and persistent Lempel-Ziv (LZ) compression, to improve application performance over the WAN. For organizations with limited backup or replication time frames but ever-increasing amounts of data to protect and replicate to remote disaster-recovery sites, Cisco WAAS can help transfer more data between data centers within a shorter period of time by increasing the utilization of the WAN connection and eliminating the transfer of redundant data across the WAN link.

Data center storage replication applications, such as EMC Recover Point, differ in nature from many other TCP-based applications due to their mode of operation and deployment environment. Data center storage replication applications are usually deployed over high-speed links and transfer large sets of data across the WAN using a limited number of TCP sessions. As a result, these applications face unique challenges in using the available bandwidth and achieving satisfactory performance.

To address the challenges of replicating data traffic over the WAN, Cisco WAAS has been tuned for storage replication applications. Cisco WAAS Release 4.0.19 is especially designed to accelerate replication tasks between data centers. This release introduces a new, dedicated device **Replication Accelerator** mode. The Cisco WAAS **Replication Accelerator** mode is optimized for high-speed links and is tuned to better utilize memory and disk resources during replication optimization. The **Replication Accelerator** mode is supported on the Cisco WAE-7341 and WAE-7371 Wide Area Application Engine (WAAE) high-end appliances.

Cisco WAAS Release 4.1

With Cisco WAAS Release 4.1, WAAS has expanded its role of WAN optimization to the branch IT services delivery platform, providing the following new benefits:

- Ease-of-deployment for initial use, and streamlined management for ongoing operations
- Improved user experience with new application-specific acceleration of:
 - Common Internet File System (CIFS)
 - Microsoft Outlook Messaging API (MAPI)
 - HTTP applications such as Oracle, SAP, and Microsoft SharePoint
 - Windows Print
 - Unix Network File Services (NFS)
- Acceleration of SSL-encrypted traffic integrates into existing security trust models
- Wide-scale delivery of live video by eliminating bandwidth upgrades and complex configuration
- Flexible delivery of branch IT services using local hosting while minimizing device footprint



Note

For more information on these enhanced services, see the *Cisco Wide Area Application Services (WAAS) V4.1 Technical Overview* at the following URL:

http://www.cisco.com/en/US/products/ps6870/products_white_paper0900aecd8051d5b2.shtml

Cisco ASA

The Cisco ASA 5500 Series includes the Cisco ASA 5505, 5510, 5520, 5540, 5550, and 5580 Adaptive Security Appliances—purpose-built, high-performance security solutions that take advantage of Cisco expertise in developing industry-leading, award-winning security and VPN solutions. The series builds upon proven technologies from Cisco PIX® 500 Series Security Appliances, Cisco IPS 4200 Series Sensors, and Cisco VPN 3000 Series Concentrators. Designed as a key component of the Cisco Self-Defending Network, the Cisco ASA 5500 Series provides proactive threat-defense that stops attacks before they spread through the network, controls network activity and application traffic, and delivers flexible VPN connectivity. The result is a powerful multifunction network security appliance family that provides the security breadth and depth for protecting small and medium-sized business (SMB), enterprise, and service provider networks while reducing the overall deployment and operations costs and complexities associated with providing this new level of security.

Through its unique Modular Policy Framework (MPF), the Cisco ASA 5500 Series brings a new level of security and policy control to applications and networks. MPF allows businesses to adapt and extend the profile of the Cisco ASA 5500 Series through highly customizable, flow-specific security policies tailored to application requirements while providing performance and extensibility through user-installable SSMs. This adaptable architecture enables businesses to rapidly deploy security services when and where they are needed, such as tailoring inspection techniques to specific application and user needs or adding additional intrusion prevention and content security such as those delivered by the Adaptive Inspection and Prevention (AIP) and Content Security and Control (CSC) SSM. Furthermore, the modular hardware architecture of the Cisco ASA 5500 Series along with flexible MPF enables the integration of future network and security, extending the outstanding investment protection provided by the Cisco ASA 5500 Series, and allowing businesses to adapt their network defenses to new threats as they arise.

For the purpose of validating this solution, the ASA 5540 was used. The Cisco ASA 5540 Adaptive Security Appliance delivers high-performance, high-density security services with Active/Active high availability and Gigabit Ethernet connectivity for medium-sized and large enterprise and service-provider networks, in a reliable, modular appliance. With four Gigabit Ethernet interfaces and support for up to 100 VLANs, businesses can use the Cisco ASA 5540 to segment their network into numerous zones for improved security. The Cisco ASA 5540 Adaptive Security Appliance scales with businesses as their network security requirements grow, delivering exceptional investment protection and services scalability. The advanced network and application-layer security services and content security defenses provided by the Cisco ASA 5540 Adaptive Security Appliance can be extended by deploying the AIP SSM for high-performance intrusion prevention and worm mitigation.

Cisco IronPort C Series Secure Email Appliance

The Cisco IronPort Secure Email Appliance protect Microsoft Exchange servers and messaging users from virus attacks, spam, and other malware before the email message is injected into the corporate network by ensuring that unwanted email messages are automatically filtered without user intervention. They significantly improve network efficiency by accurately blocking up to 80 percent of incoming spam and malware at the connection level.

The Cisco IronPort C Series contains a powerful multilayered approach to email security, providing advanced threat prevention, blocking spam and viruses, and enabling corporate data loss prevention and remediation. Key features include:

- Spam protection
 - Cisco IronPort Senderbase Reputation Filtering typically rejects more than 80 percent of all incoming connections that are from senders known through Senderbase to be malicious. Senderbase, as with all Cisco IronPort technologies, strives for and maintains the lowest false-positive rates in the industry. Rejecting connections from known malicious senders is highly effective in combating spam, and greatly increases the effective performance of the platform.
 - Cisco IronPort antispam provides best-in-class spam protection with the lowest false-positive rate in the industry while maintaining very high spam catch rates; Cisco IronPort antispam uses a suite of cutting-edge technologies and strategies that consistently outperform competitive solutions.
- Virus protection
 - Cisco IronPort security appliances provide a choice of anti-virus vendor, all of which provide effective anti-virus protection.
- Data-loss prevention and business-class email messaging

- Cisco IronPort Business-Class Email (BCE) provides a cutting-edge solution to on-demand and policy-based email encryption. BCE uses Cisco IronPort PXE encryption, which is based on strong industry-standard encryption technologies, supports envelope-based use policies such as secure forward and secure reply, and provides guaranteed delivery and read receipts.
- Cisco IronPort DLP augments IronPort BCE with a growing set of rules and policy templates to detect and manage outbound content, helping to protect your organization from unintended data and intellectual property loss.

The Cisco IronPort C Series is also used as a full replacement of the Microsoft Exchange 2007 Edge Transport role for SMTP SmartHost functionality. It will be connected to the Exchange 2007 Hub Transport servers for inbound/outbound SMTP message routing. They reside at the edge of the network behind an ASA firewall.

Cisco MDS

The Cisco MDS 9500 Series and 9200 Series supports network-hosted storage services that can be extended to any SAN-connected host or system. By providing open APIs for Cisco and third-party applications such as EMC and SANTap, the Cisco storage networking platforms allow storage services to be deployed transparently. This approach enables scalable performance, and reliable service to support changing customer needs.

Cisco MDS 9500 Series Multilayer Directors share a common architecture, the Cisco MDS 9000 SAN-OS operating system, and switching and services modules that are backward and forward compatible throughout the entire Cisco MDS 9500 Series Multilayer Directors and 9200 Series Multilayer Fabric Switches. Support for RecoverPoint with SANTap services housed on a Storage Services Module is identical for a Cisco MDS 9500 Series or a Cisco MDS 9200 Series.

This project used Cisco MDS 9500 director class switches in the primary data center to host storage access from the ESX servers running Exchange and related services. In addition, these directors tied into RecoverPoint, DMX storage arrays and Nexus 500 switches. Speeds of 2Gb FibreChannel and 4Gb FiberChannel were used to connect storage and hosts. These MDS 9509s were populated with dual supervisors to provided reliability in the storage network to match the reliability exhibited by clustering VMware ESX servers and the highly available EMC DMX Array.

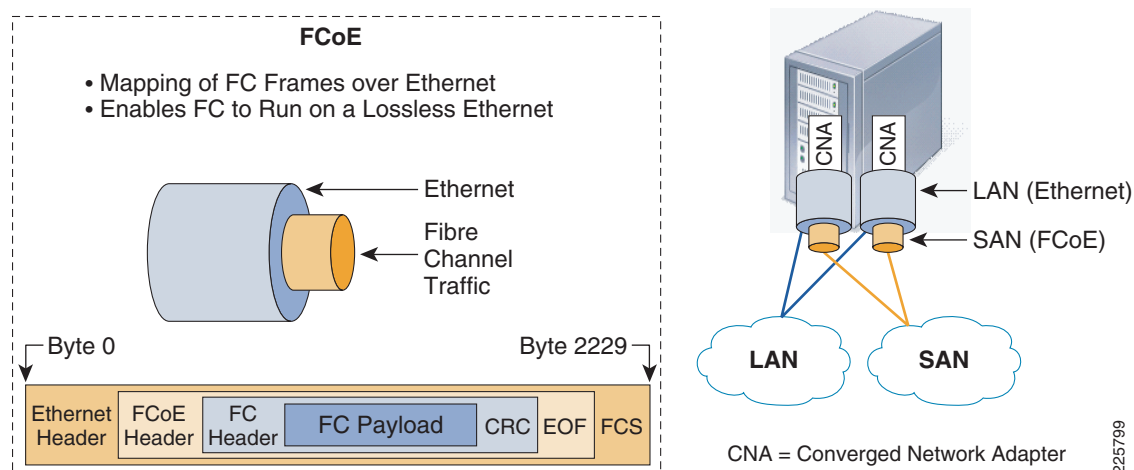
The recovery data center used the Cisco MDS 9222i with 18 integrated 4Gb FibreChannel ports and an additional Storage Services Module (SSM). The Cisco MDS 9200 Series delivers state-of-the-art multiprotocol and distributed multiservice convergence, offering high-performance storage area network (SAN) extension and disaster recovery solutions, intelligent fabric services, and cost-effective multiprotocol connectivity. The Cisco MDS 9222i is used as a more economical route for a smaller recovery data center. It does not have the high availability that is exhibited in the larger director class Cisco MDS 9500 Series, but it has a modular slot that can be used for an SSM. This modularity and rich feature set of the Cisco MDS 9200 makes this a very flexible switch capable of supporting integrated SANTap with EMC RecoverPoint.

The Cisco MDS 9222is connected the MDS storage fabric to the Nexus 5000s Unified Fabric with dual 4 Gb FibreChannel connections in a port channel connected to the Nexus 5000s FibreChannel expansion module. ESX servers and corresponding DMX storage was connected to the MDS Ports at 2Gb and 4Gb FibreChannel.

Cisco Nexus 5000

Unified Fabric helps to consolidate separate Local Area Network (LAN), Storage Area Network (SAN) and server cluster network environments into a single unified network. Unified I/O at the port and cable level is an important benefit of a unified fabric. The transition from today's multiple heterogeneous data center networks to the unified data center network of tomorrow will be an evolution that will happen in phases. Cisco is helping lead this transition in the standards organizations, working to define the protocol suite that will be the foundation for the new unified data center. An important pillar of this foundation is FibreChannel-over-Ethernet (FCoE). FCoE allows FibreChannel frames to be encapsulated in Ethernet packets without using TCP/IP and is one of the technologies enabling unified I/O. [Figure 3](#) shows the encapsulation of FCoE.

Figure 3 FCoE Quick Overview



Connectivity to the Nexus 5000 is facilitated by Converged Network Adapters (CNAs). Several ESX hosts in both data centers in the Exchange test environment have CNAs installed providing 10Gb converged I/O.

FCoE is required to be a lossless Ethernet network, with switching devices that have internal architectures designed to offer a no-drop packet capability and network flow control mechanisms to enable lossless transmission of packets across the Ethernet infrastructure. FCoE provides several advantages over existing approaches to I/O consolidation:

- Forward compatibility with existing FibreChannel SANs by preserving well-known FibreChannel concepts such as virtual SANs (vSANs), World Wide Names (WWNs), FibreChannel IDs (FCIDs), and zoning to servers and storage arrays.
- A high level of performance, comparable to the performance of current Ethernet and FibreChannel networks, achieved by using a hardware-based Ethernet network infrastructure that is not limited by the overhead of higher-layer TCP/IP protocols.
- Exceptional scalability of Ethernet at the highest available speeds (1, 10, and eventually 40 and even 100 Gigabit Ethernet).
- Gatewayless technology, avoiding the overhead of higher-layer TCP/IP protocols and simplifying operations and management (no change to the management infrastructure currently deployed in SANs).

VMware Technology Overview

This section discusses the VMware technologies used in the design presented in this document. The technologies used are as follows:

- VMware Infrastructure 3
 - VMware ESX
 - VMware High Availability (HA)
 - VMware DRS
 - VMware VMotion
- VMware vCenter Site Recovery Manager

VMware ESX

By taking advantage of the inherent benefits of a virtualization-based platform, an Exchange deployment using VMware Infrastructure with ESX hosts offers a variety of availability options. Each of these options provides different levels of both protection and cost, capable of meeting the unique high-availability requirements of any Exchange environment. Many tools are available from EMC and Cisco to facilitate both in-site and remote-site availability and recovery.

An important benefit of virtualization is abstraction of the operating system and application from the underlying physical server hardware. This abstraction is extremely useful in disaster-recovery scenarios because it eliminates the traditional requirement of physical server-based disaster recovery to provide identical hardware at the disaster-recovery site. Any virtual machine can be brought online on any supported ESX server without worrying about hardware or software compatibility. The ability to run multiple virtual machines on a single server also reduces the costs of a disaster-recovery solution through consolidation of Exchange components and services on fewer physical servers than would normally be required. Thus, all the necessary Exchange server roles and Active Directory components can be deployed in virtual machines at a disaster-recovery site with minimal hardware for speed recovery in a disaster situation.

Finally, virtual-machine encapsulation means that an entire Microsoft Exchange 2007 server can be contained in a small set of files, simplifying replication to disaster-recovery sites. You can move an entire virtual machine with a simple file copy.

**Note**

It is important to understand Microsoft's licensing support for operating system and application deployment in virtual machines. For more information, refer to the following URL:

<http://support.microsoft.com/?kbid=897615>

VMware High Availability (HA)

VMware HA automatically restarts virtual machines that run on hosts that experience a failure; for example, if a motherboard fails or the host panics. This solution provides simple, low-cost protection for virtual machines by guarding them against physical host failure. If a server hardware outage occurs, VMware HA automatically restarts all virtual machines on another VMware ESX server, minimizing disruption to the Exchange environment. Simple to set up, VMware HA protects every virtual machine without requiring complex clustering software.

VMware DRS

VMware Distributed Resource Scheduler (DRS) tracks the performance of virtual machines and, depending on the configuration, recommends target hosts for best performance or actually migrates hosts based on policy. With VMware DRS, virtual machines are dynamically load balanced across an entire pool of server resources. DRS collects resource usage information for all hosts and virtual machines and generates recommendations for virtual-machine placement. You can apply these recommendations manually or automatically. DRS can dynamically load-balance all virtual machines in the environment by shifting workloads across the entire pool of ESX servers, ensuring that critical Exchange virtual machines in the environment always have the CPU and RAM resources they need to maintain optimal performance.

VMware VMotion

VMware VMotion allows clients to move running virtual machines from one physical host to another with no effect on end users. VMotion can easily move under- or overused Microsoft Exchange servers to meet the increasing demands of businesses and end users.

VMware vCenter Site Recovery Manager

The solution is designed around VMware Site Recovery Manager to accelerate recovery and ensure successful recovery of a Microsoft Exchange Server by automating the recovery process and eliminating the complexity of managing and testing recovery plans. VMware Site Recovery Manager enables this solution to provide disaster recovery for Microsoft Exchange that is rapid, reliable, and manageable so that businesses can meet recovery objectives. By eliminating complex manual recovery steps and enabling nondisruptive testing of recovery plans, Site Recovery Manager removes the risk and worry from disaster recovery, helping businesses protect all of their important systems and applications.

- SRM uses block based replication with Storage Recovery Adapters installed on the SRM Server. This integration of hardware and software supports the most demanding application business continuance needs, in this case, a failover following a disaster. Multiple storage adapters can be supported by SRM depending on which storage array you have installed. SRM fully integrates into VMware vCenter Server and allows you to create and manage recovery plans from one common interface. Integration into vCenter allows discovery and display of virtual machines protected by storage replication using integrations certified by storage vendors. Extend recovery plans with custom scripts, store, view and export results of test and failover execution from VMware vCenter Server.
- SRM provides for isolated network testing of recovery plans with real replicated data. Customization is supported with integration into ESX and vCenter for IP address changes. With identical interfaces for testing and recovery, SRM provides for recovery plan testing and execution with a single button. Full monitoring of SRM actions is available in addition to monitoring tasks and events on specific ESX servers in vCenter.

Microsoft Exchange Server 2007 Overview

The Microsoft Exchange Server 2007 offers many advantages to customers in the form of built-in protection, flexible access methods, and operational efficiency. Customers are looking for ways to cut cost and increase productivity while ensuring that there is high availability. Microsoft Exchange Server 2007 was designed to offer solutions to these most demanding customer messaging requirements and do so for a variety of endpoints, from any location and to provide access to messaging resources in a secure and highly available manner.

Some of these customer requirements are met by enabling the following:

- Integrated message filtering
- Business continuance through several clustering and disaster recovery options
- Endpoint security for a variety of access methods which include a web client, Outlook, mobile, and POP/IMAP
- Flexible policy creation, management and reporting for legal compliance needs
- Streamlined setup, administration and management via the Microsoft Exchange Management Console, Exchange Management Shell, and Systems Center products
- Scalability and performance improvements through a x64-based architecture, increased memory support, and more intelligent message routing

There are many feature improvement and advantages of using Microsoft Exchange Server 2007 as well as comparisons with Microsoft Exchange Server 2003. Additional information on these features, advantage, and comparisons can be found at the following URL:

<http://www.microsoft.com/exchange/evaluation/default.aspx>

Microsoft Exchange Server 2007 requires an existing Microsoft Active Directory (AD) deployment and leverages AD as a means to store and share information within the Exchange environment. More information regarding the planning and deployment of Microsoft Active Directory in support of Exchange Server 2007 can be found here: <http://technet.microsoft.com/en-us/library/bb123715.aspx>.

**Note**

All references to Exchange Server 2007 used in testing imply the most up-to-date version of Exchange at time of validation, which is Exchange Server 2007 Service Pack 1 (SP1).

Microsoft Exchange 2007 Server Roles

There are five roles in Microsoft Exchange Server 2007. Each role serves a unique purpose within the Microsoft Exchange architecture and is flexible enough to be deployed in various sized organizations with varying requirements.

Most roles can be installed together on a single platform or can be deployed completely independent of one another. Small-medium customers can leverage the diverse number of Microsoft Exchange Server 2007 features while limiting the amount of hardware required for deployment by deploying the roles on the same server. Large organizations can leverage having multiple roles deployed in a redundant fashion on independent hardware platforms in geographically dispersed locations.

The five roles in Microsoft Exchange Server 2007 are:

- Client Access Server (CAS)
- Hub Transport (HT)
- Mailbox Server (MBX)

- Edge Transport (ET)
- Unified Messaging (UM)

The following sections will describe three of the five roles at a high-level and is not meant to be a full tutorial on the architecture, design, and operation of each role. The UM and ET roles were the only roles that were not tested in this multisite data center design.


Note

In this document, the ET role has been replaced by the Cisco IronPort C-Series Secure Email Appliance. The IronPort Secure Email Appliance provides SMTP SmartHost functionality as well as industry leading SPAM, virus, policy enforcement, and email encryption.

Detailed information on the Microsoft Exchange Server 2007 product, architecture, and design is found at the following: <http://www.microsoft.com/exchange> or <http://technet.microsoft.com/en-us/library/bb124558.aspx>

Client Access Server

The client access server (CAS) provides access for a variety of client endpoints. The CAS role was formerly known as the Exchange frontend server. The CAS role supports access through the following methods:

- Microsoft Outlook Web Access (OWA)
- Post Office Protocol Version 3 (POP3)
- Internet Message Access Protocol Version 4 (IMAP4)
- Microsoft Exchange ActiveSync client
- Microsoft Outlook Anywhere

The CAS role also supports various other web services such as the offline address book (OAB) distribution and the autodiscover service. The list above shows that the CAS role can provide access to messaging services through many different endpoint types such as computers with web browsers, Outlook outside of the corporate firewall, email clients using POP3/IMAP4 and even mobile devices. Endpoints using another method of access such as Messaging Application Programming Interface (MAPI) most often connect directly to the mailbox server (MBX) role while within the corporate firewall (see [Mailbox Server, page 22](#)).

In the simplest terms, the CAS role provides a frontend service for the MBX role for non-MAPI connections. The CAS communicates directly with the MBX. The CAS role is optional, if there are no requirements to use non-MAPI clients.

Microsoft recommends the deployment of multiple CAS for performance, scalability, and availability purposes. The Microsoft Exchange Server 2007 fully supports multiple CAS role servers to be active simultaneously. This is ideal for an active/active multisite data center design.

Hub Transport Server

The Hub Transport (HT) role, formerly known as the bridgehead server, is the central role for intelligent message routing delivery and policy control. Unlike the CAS and Edge Transport (ET) roles, the HT is required.

All mail flow external to the organization and internal within the organization is handled by the HT role. The HT role can use the ET as an SMTP relay for messages going to/from the Internet or it can handle the SMTP relay role on its own. The HT communicates directly with the MBX, other HT roles, and the ET.

Messaging routing within the Microsoft Exchange environment requires the configuration of Active Directory (AD). AD is used to ensure that optimal message routing is accomplished within and between AD sites. This is quite different from previous Microsoft Exchange versions where routing groups were the primary method for messaging routing.

As was the case with the CAS role, it is recommended by Microsoft to deploy multiple HT roles for performance, scalability and availability purposes. Microsoft Exchange Server 2007 fully supports for the HT role to have multiple servers active simultaneously. This is ideal for an active/active multisite data center design.

Mailbox Server

The mailbox server (MBX) role is the database for all user messaging data. Users are homed to a particular MBX and associated storage group. As mentioned before, MAPI-based clients such as those running Microsoft Outlook connect directly to the MBX while within the corporate firewall. The MBX role is a required component of an Exchange Server 2007 deployment.

Microsoft Exchange Server 2007 has several options for maintaining high availability (HA) of the MBX role to include Local Continuous Replication (LCR), Cluster Continuous Replication (CCR), Standby Continuous Replication (SCR – Service Pack 1-only) and Single Copy Cluster (SCC). For more information from Microsoft on these solutions, refer to the following URL:

<http://technet.microsoft.com/en-us/library/bb124721.aspx>.

Also, Cisco has validated CCR and SCR in a Cisco multisite data center design and the details can be found at the following URL:

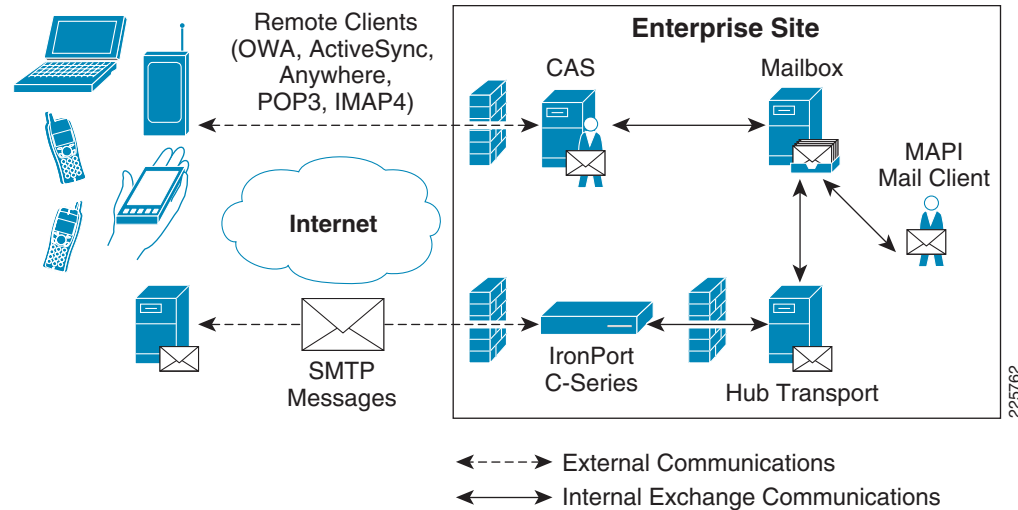
<http://www.cisco.com/en/US/docs/solutions/Verticals/mstdcmsftex.html>

In this document, HA/DR of the MBX role will be provided using technologies from Cisco, EMC, and VMware. Specifically, the MBX role and associated storage will be replicated using EMC RecoverPoint, Cisco MDS/SSM/SANtap and VMware SRM to the secondary data center site.

The MBX role is the only Exchange Server 2007 role that does not support an active/active configuration. However, the MBX role is also the only role that supports clusters. Therefore, more than one MBX can be deployed for scalability and availability purposes, but a user can only be connected to a single MBX that user is associated with.

The MBX communicates directly with the CAS and HT.

Figure 4 shows a high-level view of the four tested Microsoft Exchange 2007 Server roles and a basic traffic flow between each role.

Figure 4 High-level view of Microsoft Exchange Server 2007 roles + IronPort C-Series

Microsoft Active Directory and Multisite Data Centers

As mentioned before, Microsoft Active Directory plays a critical and required role in the Microsoft Exchange Server 2007 environment. In the testing conducted, there were two AD deployment options that were validated. The first was using a single AD site for two active data center locations and the second was using an AD site for each data center location by using the Microsoft Active Directory Sites and Services capability to create and manage AD replication between sites.



Note

All designs and references in this document are based on using Microsoft Windows Server 2008. Microsoft Exchange Server 2007 with SP1 is required to support Microsoft Windows Server 2008.

Single AD Site — Multiple Data Center Locations

There are many things to consider in a “normal” AD deployment model that will determine the success or failure of a scalable and available AD implementation. Adding the additional issues involved with now spanning a single AD site to multiple physical locations that can be geographically dispersed by great distance may be too great for many customers to undergo. Some, but certainly not all, of the considerations that a customer needs to account for are:

- Available network bandwidth and latency between each data center
- Suitable AD replication schedule between domain controllers/global catalog servers
- Contention between AD replication and other application/network traffic between data centers
- Containment of AD objects to a local site for management and security purposes

The considerations listed above will most often dictate that the data centers are close enough to each other to provide adequate bandwidth and low latency.



Note

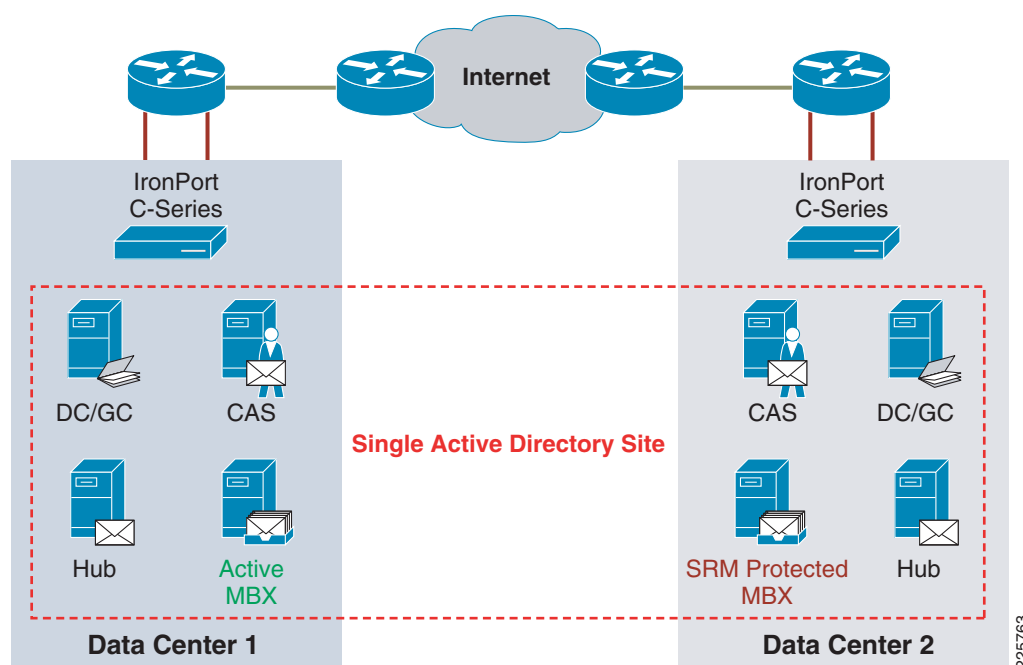
This document is not intended to provide the required knowledge for AD planning and implementation for Microsoft Exchange Server 2007. Additional information related to AD requirements for Exchange Server 2007 can be found at the following URL:

<http://technet.microsoft.com/en-us/library/bb123715.aspx>.

The single AD site model was used and tested as it was the best model to allow for nearly all Microsoft Exchange Server 2007 components to function in an active/active role. As mentioned before, the mailbox server role is the only role that cannot support load balancing and/or active/active configurations. The CAS and HT roles can support an active/active data center deployment. The reader must research and understand the Microsoft AD and Exchange Server 2007 implications of such a design before considering it for deployment.

Figure 5 shows a high-level overview of the single AD site model as it resides within two active data centers. The dashed box indicates that both DC locations are within the same single AD site. The only role in this scenario that cannot support an active/active configuration is the mailbox server role. In this example, the MBX role is protected by EMC RecoverPoint and VMware SRM and is replicated to Data Center 2. In the event of a failure at Data Center 1, SRM activates the MBX role at Data Center 2. The only thing that needs to be updated is the IP address and DNS entry; for more details, refer to the configuration and design of Recover Point and SRM in [Customization, page 80](#). All other roles shown can be active in both DC locations, simultaneously.

Figure 5 *Single Active Directory Site—Two Data Center Locations*



Multiple AD Sites—Multiple Data Centers

While the single AD site model allows for the ability to have most Exchange Server 2007 roles in an active/active configuration, the requirements for supporting such a design may outweigh the advantages. As discussed in [Single AD Site — Multiple Data Center Locations, page 23](#).

There are many considerations to plan for when dealing with a single AD site model for Exchange Server 2007. The AD, Exchange, and network administrators must balance the active use of resources in all data center locations against the management and cost associated with the support of full active-use of each resource in each location.

The model of supporting at least one AD site per data center location is easier to plan and deploy as well as support, especially when the data centers are geographically dispersed. If the primary goal is that of site-wide disaster recovery versus load balancing between sites, the multiple AD site model is more appropriate. With that said, it is possible to have some roles of the Exchange Server 2007 deployment be active/active in the multiple AD site model. One example of using an active/active configuration with multiple AD sites is with the CAS Deployment. A deployment may have multiple CAS roles per DC location and each DC has one or more AD sites. If site load balancing directs a user request to a CAS role located in a different DC (subsequently a different AD site) than the user belongs to, a feature known as CAS-CAS proxying or, depending on the setup, CAS-CAS redirect can still connect the user to the correct CAS role for their site which then connects to the correct mailbox server. This feature allows for the CAS roles to be active at both DC locations. More information can be found on CAS-CAS proxying at: <http://technet.microsoft.com/en-us/library/bb310763.aspx>.

Similar types of considerations exist for both single AD and multiple AD site models but are less stringent for the multiple AD site model. Microsoft Active Directory Sites and Services is designed to implement and deploy multiple AD sites, their resources and schedules for AD replication. As they apply to AD and Exchange, network bandwidth and latency requirements for the multi-AD site design are also less stringent because the links between DC locations are mostly used for AD replication versus full-time use for AD and active/active mail traffic flow.

Tested Microsoft Exchange Server 2007 Deployment Models

Microsoft Exchange Server 2007 Layout

There are many possible combinations of Exchange Server 2007 implementations. In this document, one implementation example is explored in more depth and has specific Cisco, EMC, and VMware product, feature, and design elements associated with the implementation example. The Exchange Server 2007 implementation example discussed in this document is based on a single Active Directory Site configuration. This example was selected as it offered a way to provide active/active for every Exchange 2007 role, except the MBX (does not support active/active).

Single-Site AD With Active/Active Traffic Flow

As discussed earlier, the goal of the single AD site design is to support an active/active data center design for Microsoft Exchange Server 2007. Having the Exchange roles in a single logical AD site eliminates the complexity and delay of having to perform an AD “fix up” on Exchange roles in the event of a site failure at the primary site. Since each Exchange role is within a single AD site, nothing within AD has to be done in the event of failure at either site to allow Exchange to continue operating.

The AD layout of this design is discussed in the previous section and illustrated in [Figure 5](#). The following section is more focused on the Exchange Server 2007 roles, their locations within the two data centers, and specific Exchange considerations for supporting the design.

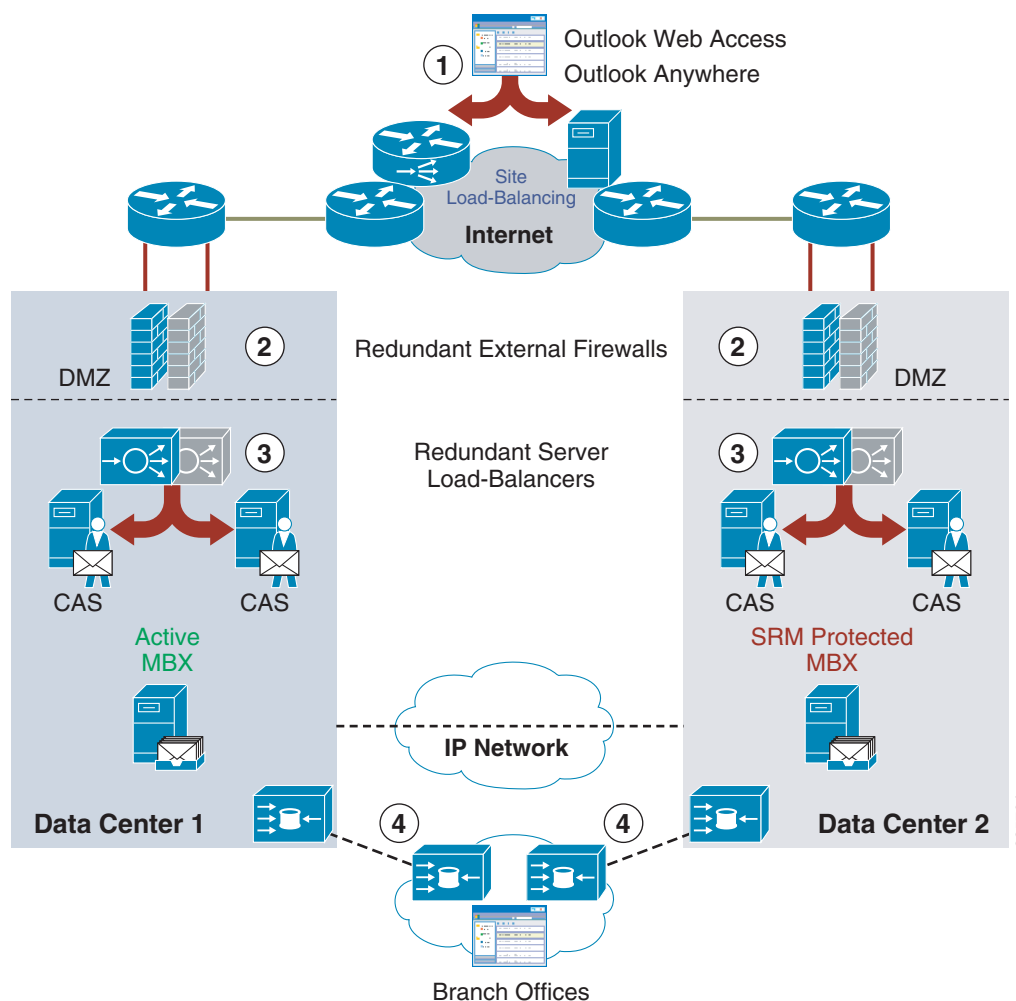
Client Access Server—Active/Active DC

Multiple Microsoft Exchange 2007 servers running the CAS role are used not only to provide fault tolerance for individual server failures and scalability to support larger volumes of sessions, but also to provide a means for supporting local site load balancing as well as geographical load balancing between sites.

In addition to being an ideal candidate for server and site load balancing, the CAS role can additionally take advantage of network optimization services and SSL-offloading.

In [Figure 6](#), a total of four Exchange servers running the CAS role are deployed at the two DC locations. In this example, the CAS role has been deployed at the data center access layer. Optionally, the CAS can be deployed at the Internet DC (IDC) edge in a DMZ context that is specifically configured for the CAS role and services both internal and external client connections. It is important to understand Microsoft does not support the placement of a firewall between the CAS and MBX roles. The dynamic port ranges used for the CAS-to-MBX connection makes it challenging to nail down the exact port-based security configuration. More information on securing data path security for each Exchange 2007 role can be found at the following URL <http://technet.microsoft.com/en-us/library/bb331973.aspx> as well as in [Security Considerations for the CAS Role](#), page 59.

Figure 6 CAS Deployment – Active/Active Data Center



The numbered objects in [Figure 6](#) correspond to the areas where the CAS role can interoperate with networking services.

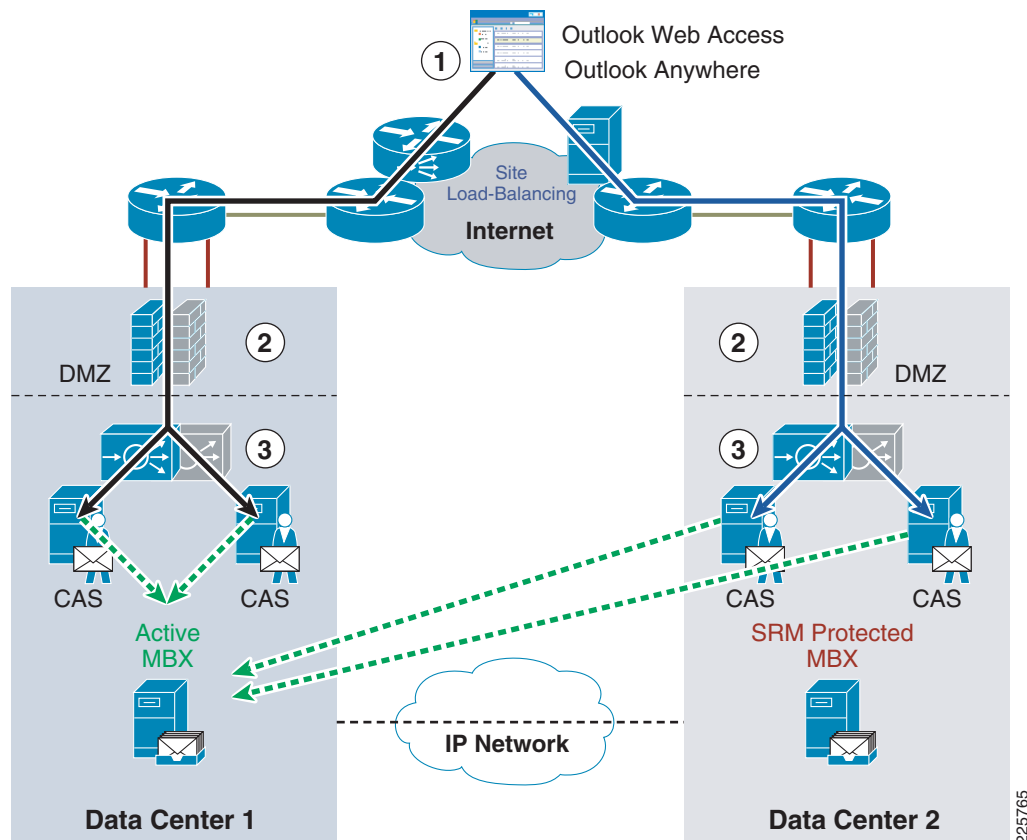
1. Site selection and load balancing for each of the CAS Web (OWA, Outlook Anywhere, Autodiscover, etc.) and non-Web (POP3/IMAP4) services through the Cisco GSS product or generic DNS round-robin.
2. The Cisco ASA or FWSM is used to provide firewall services.

3. The Cisco ACE module can be deployed for Layer 4 through Layer 7 load balancing and can monitor the health of the CAS services and intelligently balance traffic amongst multiple CAS roles as well as report availability to the Cisco GSS. Also, at the same location, SSL-offload can be performed on the CAS role to help scale services such as OWA which uses HTTPS. The SSL-offload features of the Cisco ACE can help reduce CPU utilization on the CAS role by offloading the encryption/decryption process for each individual HTTPS session.
4. If branch office users connect to the CAS services located at either of the active data center locations, the Cisco WAE product can perform WAN optimization on the sessions to reduce bandwidth utilization, optimize the TCP sessions, and reduce or eliminate duplicate data being transmitted between sites. It is important to note that the Microsoft Exchange and network administrators work together to understand the advantages and disadvantages of optimizing CAS services by using a WAN optimizer. A decision needs to be made on whether or not client-to-server encryption of data in transit is more or less important than network optimization of the same traffic. End-to-end encryption and optimization of the payload are generally mutually exclusive. Optimization for encrypted MAPI traffic and other SSL/TLS traffic will be available in a future release of WAAS slated for early CY 2009.

Some customers will disable encryption for the OWA, Outlook Anywhere, or Outlook MAPI connections and rely on site-to-site VPN for encryption between the branch and DC. This still allows for encryption across the WAN while enabling optimization of the MAPI or HTTP flows.

Figure 7 shows an example traffic flow for the CAS role when an external user is connecting to CAS services, in this case OWA.

Figure 7 CAS Traffic Flow Example



In [Figure 7](#), the following occurs:

-
- Step 1** The Cisco GSS is used to intelligently select the most appropriate site to load-balance the OWA traffic to. Alternatively, DNS round-robin can be used but DNS round-robin does not offer advanced features such as proximity-based routing, VIP tracking and service availability for each site and resource. The solid lines indicate possible paths the session may take depending on the load-balancing decision made for the CAS.
- Step 2** Once the OWA traffic has been routed to a specific DC from Step 1, additional network services can be leveraged for security, performance and availability. Firewall services that are located in the DMZ can be applied to the CAS-targeted flows.



Note It is common for customers to deploy the Cisco ASA in front of (in series) or beside (in parallel to) the Microsoft ISA 2006 product so that there is one tier of firewall protection by the ASA and another tier of protection via the OWA publishing capability in the Microsoft ISA 2006 product. In a Cisco ASA + Microsoft ISA 2006 configuration, only traffic destined to the CAS addresses or Cisco ACE VIP address for the CAS is directed to the ISA.

- Step 3** Intelligent server load balancing and SSL-offload can be performed on the OWA services at this step. After a specific server running the CAS role has been selected and the above mentioned services performed, the OWA session is successfully connected to the CAS. The dotted lines indicated the paths from the CAS to the appropriate mailbox server that is associated with the user mailbox object.

In this example, all CAS-to-mailbox connections terminate on the active MBX located in DC1. In the event of a node failure at DC1, all CAS-to-mailbox connections automatically (or manually) terminate at the now activated MBX that was replicated by EMC RecoverPoint and recovered through VMware SRM at DC2.

Hub Transport Server—Active/Active DC

The Hub Transport (HT) role has a built-in load balancing (round-robin) capability that allows each HT role to select another Microsoft Exchange 2007 role to route mail to. This is true of mail routing internal to the same AD site, a different AD site, and for external SMTP relay to IronPort secure email appliances (via Send and Receive Connectors).

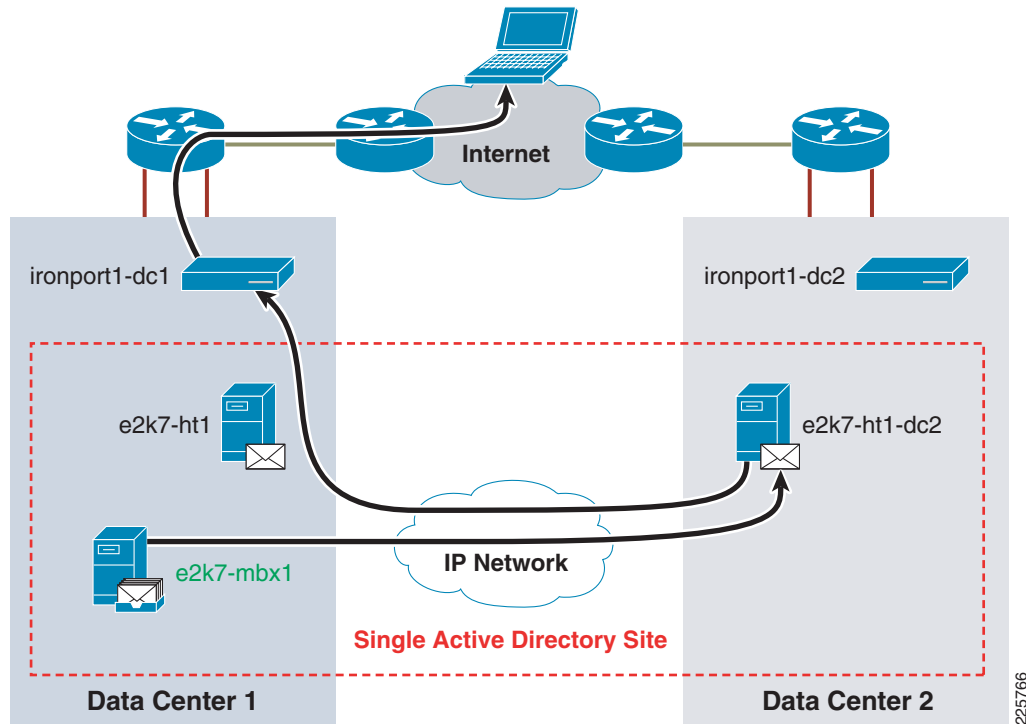
Other than providing firewall services for communication to/from HT roles in the active/active design, no other significant network services are being discussed in this document. Communications between HT roles and other roles is usually encrypted using Transport Layer Security (TLS) and also load-balanced within the HT role itself. External load balancing, SSL-offload and network optimization do not offer the same level of advantages for the HT role as with the CAS and ET roles.

One thing to note regarding the HT role and the single AD site design being discussed in this section is the communication between the HT and the mailbox server role. By default, mailbox servers use a list of known HT roles within their own AD site and will make a round-robin based decision on which HT role to send mail traffic to. Since the single AD site model spans multiple DC locations, the mailbox server may load-balance mail traffic to HT roles at a different DC location only to have the HT make its own load-balance decision and send the mail back to another Exchange server role in the original DC location.

[Figure 8](#) illustrates a likely scenario where mail from a particular mailbox server needs to be routed to an external mail recipient. The mailbox server, by default, will load-balance to all HT roles in its own AD site (in this example, a single AD site spans both DC1 and DC2). In this example, a message needs to be routed from the mailbox server **e2k7-mbx1** to an external recipient and the round-robin choice this

time is to **e2k7-ht1-dc2** located in DC2. The message traverses the DC-to-DC interconnect link and arrives at **e2k7-ht1-dc2**. As mentioned before, the HT roles also do round-robin load balancing on Send Connectors to the IronPort appliance and the round-robin choice this time is **ironport1-dc1** in DC1. Again, the message traverses back across the DC-to-DC interconnect and arrives at **ironport1-dc1** and subsequently is relayed to the Internet.

Figure 8 Sub-optimal Active/Active Mail Flow Example



If this message routing is of concern to the network and Exchange administrator, it is possible to force the mailbox server to use only those HT roles that are physically located in the same DC as the mailbox server. This is done by modifying the **SubmissionServerOverrideList** for the mailbox server and list out only those HT roles located in the same DC as the mailbox server. For example, if the administrator wanted the mailbox **e2k7-mbx1** to only use **e2k7-ht1** because it was physically located in the same DC, the administrator could run the following command in the Exchange Management Shell:

```
[PS] C:\>Set-MailboxServer -Id:e2k7-mbx1 -SubmissionServerOverrideList: e2k7-ht1
```

The setting is now altered in the mailbox server properties in Active Directory:

```
[PS] C:\>Get-MailboxServer -Id:e2k7-mbx1 | fl SubmissionServerOverrideList
SubmissionServerOverrideList : {E2K7-HT1}
```

The mailbox server will now only use the configured HT roles listed. However, the administrator needs to ensure this setting is changed if there were a failure of the designated HT roles. An example of such a failure would be in the case of a total site failure in DC1. The mailbox server(s) now active in DC2 needs to be configured to use the HT roles located in that physical DC site. Remember that the MBX in this scenario is replicated by EMC RecoverPoint and VMware SRM. When SRM performs the recovery plan and activates the MBX VM it will change the IP address and do a DNS update of the MBX for DC2 but everything else about the MBX will be as it was in DC1. Without modifying the override list, the DC2 mailbox servers would only try to communicate with HT roles located in a site now unreachable.

(due to site failure). VMware SRM does allow custom commands to be executed once the MBX VM comes online. A Windows PowerShell command could be scripted to run the above command post the VM activation.

Mailbox Server—Active/Active DC

Previous sections in this document discussed the differences in how a CAS in the same AD site as the MBX connects the MBX user directly to their home MBX VM versus when a CAS in a different AD site as the MBX performs either CAS-to-CAS proxying (internal connection) or CAS-to-CAS Redirection (external URL redirection). With a single AD site implementation, all CAS roles in both DC locations can connect logged in users directly to their home MBX.

The other consideration for the MBX in this design, again, has to do with the HT role and ensuring optimal mail flow between the MBX and HT roles. The section above addressed how to deal with suboptimal mail flow for MBX-to-HT connections.

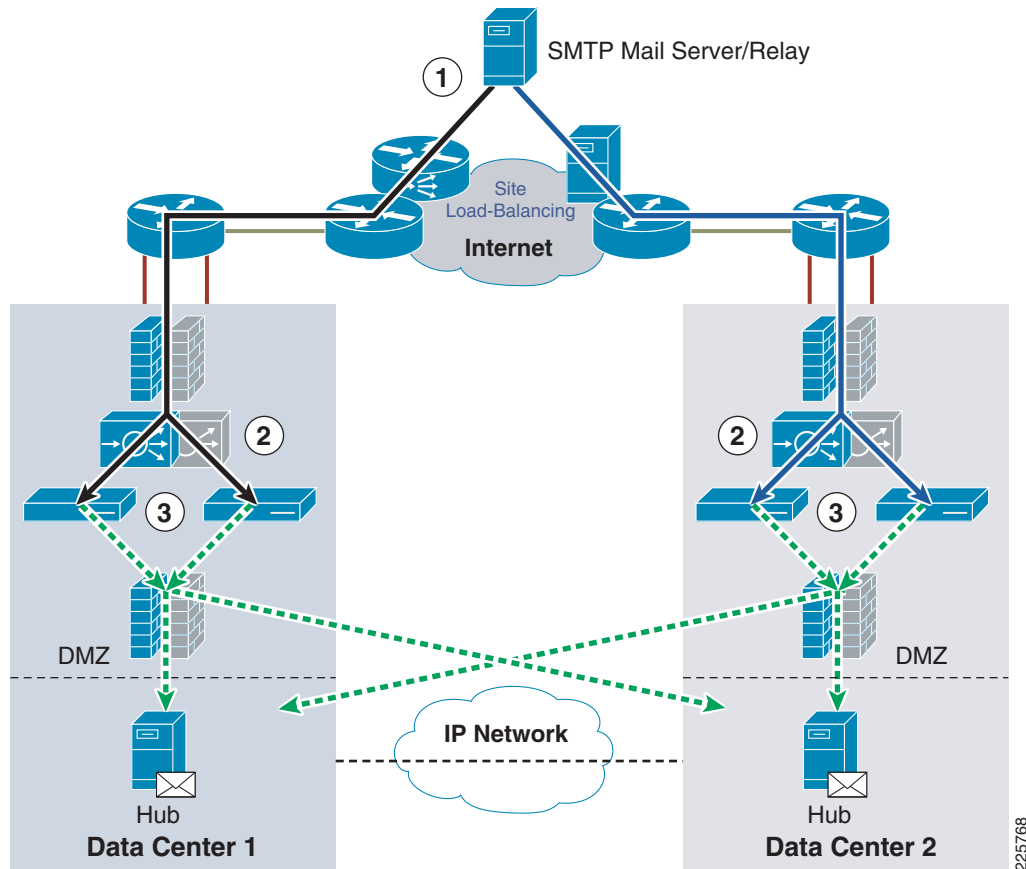
IronPort C-Series Secure Email Appliance—Active/Active DC

The IronPort Secure Email Appliance should be deployed in a redundant way and are an ideal candidate for site load balancing. Deploying service and/or site load balancing for the IronPort provides additional fault tolerance within the site and between sites as well as scalability of the IronPort services.

The IronPort is deployed at the edge of the network in a secured and load-balanced DMZ context. Some organizations will deploy their IronPort in a design that centralizes either inbound or outbound SMTP messages or perhaps both. One example of this is where an organization has multiple sites around the world and designs the Exchange environment to only allow SMTP messages into the organization via a site or sites within the United States, but permits outbound SMTP messages from any site internationally. Regardless of the reasons why or how the organization wants external SMTP messages to flow, scalability, security, and availability are always required.

[Figure 9](#) shows an example traffic flow for the IronPort email appliances when an external SMTP message is sent from the Internet to the organization.

Figure 9 *IronPort Secure Email Appliance SMTP Flow Example*



The traffic flow for [Figure 9](#) is very similar to that of the CAS role:

-
- Step 1** The Cisco GSS can be used to intelligently select the most appropriate site to load-balance the SMTP traffic to. Alternatively, DNS round-robin can be used to resolve the MX record. The solid lines indicate possible paths the session may take depending on the load-balancing decision made.
- Step 2** Once the SMTP traffic has been routed to a specific Data Center from Step 1, additional network services can be leveraged for security, performance and availability, such as firewall and server load balancing.
- Step 3** After a specific IronPort appliance has been selected and the above mentioned services performed, the SMTP messages are sent directly to an internal HT role. SMTP messages inbound and outbound between the HT and IronPort are load-balanced using a round-robin method. The dotted lines indicate possible paths that may be taken from the IronPort to the HT roles based on the round-robin selection outcome.
-

Optimization and Availability Support for Microsoft Exchange Server 2007 in a Cisco Multisite Data Center

It is important to give a brief overview of what roles within Exchange Server 2007 support various methods of load balancing (LB), fault tolerance (FT), network optimization, and SSL-offload. Some of these methods are handled within the Exchange and/or server architecture, in standard ways such as with DNS or NIC-teaming or those methods that are provided by networking components.

In [Table 1](#), the Exchange 2007 messaging service is listed along with the method used to provide LB, FT, network optimization, and SSL-offload.

Table 1 *Microsoft Exchange Server 2007 Services and LB, FT, HA Methods Supported*

Microsoft Exchange 2007 Service	Site Load-Balancing	Server Load-Balancing	Fault-Tolerance	Network Optimization	SSL-Offloading
Client Access Server	Cisco Global Site Selector (GSS) and/or DNS round-robin	Cisco ACE, Microsoft Network Load-Balancing (NLB) or DNS round-robin	NIC-teaming, multiple CAS roles	Cisco WAE	Cisco ACE
Hub Transport Role	N/A	Handled internally by Microsoft Exchange	NIC-teaming, multiple Hub Transport roles	N/A	N/A
Mailbox Server	N/A	N/A	NIC-teaming, EMC/Cisco/VMw are replicated MBX/Storage Clusters (LCR, CCR, SCR, and SCC)	Cisco WAE	N/A
IronPort Secure Email Appliance	Cisco Global Site Selector (GSS) and/or DNS round-robin	Cisco ACE, Microsoft NLB or DNS round-robin	Redundant IronPort appliances	N/A	N/A

Design and Implementation Details

Design Goals

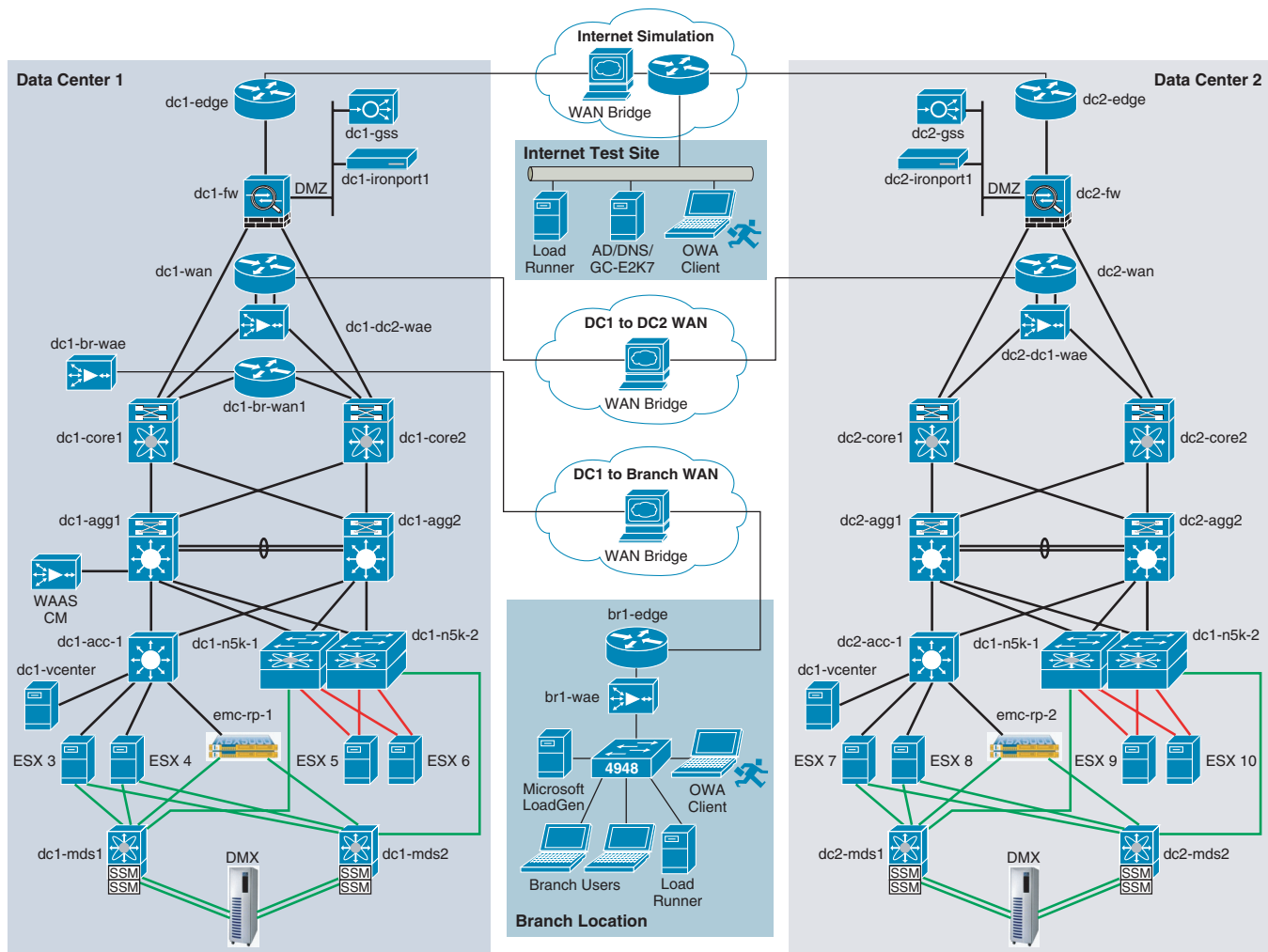
The enterprise network is a platform constructed to support a myriad of business functions; more specifically, applications. The traditional perception of the network relegates its role to one of data transport, providing a reliable fabric for the enterprise. This is a fundamental responsibility of the network infrastructure and should be enhanced rather than neglected. In addition to transport, the ubiquitous nature of the enterprise network fabric allows the introduction of intelligent network services to support business applications. This evolution of the network as an enterprise service platform is natural and supports the following application objectives:

- High availability
- Scalability
- Security
- Optimization

The Cisco data center architecture is a holistic approach that allows the network and the applications it supports to work together. The primary goals of this design are to increase the performance, availability, scalability, and manageability of enterprise applications within and between data centers, while simultaneously providing a secure environment. In addition, this design reduces the complexity and implementation time of enterprise applications in the data center using virtualization technologies and network design best practices. The remainder of this document focuses on each of these objectives when deploying a Microsoft Exchange application using the services of the Cisco, EMC, and VMware data center infrastructure and Cisco empowered branch solutions.

Enterprise Data Center Design

[Figure 10](#) represents the multisite test deployment of Microsoft Exchange 2007. Each data center site provides a highly available and robust network, storage, and application infrastructure for the local Exchange environment. The introduction of multiple data centers extends the *n*-tier Exchange application model between sites. Therefore, the network must address the state of the application-tier at each site to provide user, application, and data recovery services.

Figure 10 **Multisite Data Center Testbed**

EMC Design and Deployment

EMC RecoverPoint and Cisco MDS Integration

As [Figure 10](#) shows, storage for the ESX servers is provided by two DMX 1000s with 2Gb FiberChannel connectivity to the storage fabric. Each LUN is configured to be available from two Cisco MDS directors each on one fabric in an active/active setup.

RecoverPoint appliances are installed in a two-host cluster at each data center. The appliances are running 3.0 Service Pack 1. Connectivity is Gigabit Ethernet to the access layer switches and Gigabit Ethernet between the two data centers across the WAN.

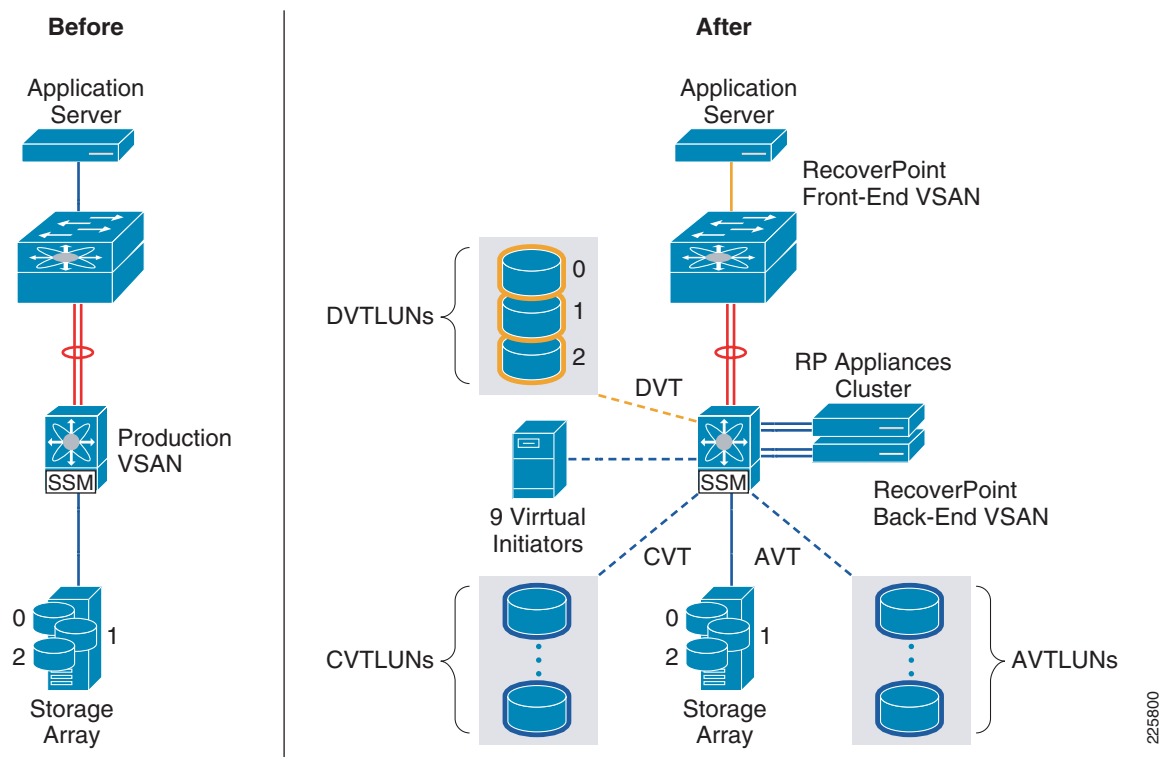
The storage network connections to the RecoverPoint appliances are 2-Gigabit FibreChannel (FC) with two connections to each fabric in the data center.

Data Center 1 consists of two MDS 9509s with Supervisor 2s running 3.2(2c) and a Switch Services Module running 3.2(3r) for SANTap services interfacing with the RecoverPoint appliances. Hosts are connected to DS-X9032 linecard at 2Gb FC. The hosts and storage are connected to the DS-X9304-18K9 at 4GB and 2Gb, respectively. In this configuration no ports are connected on the SSM ports. Each host connects to both MDSs as do the RecoverPoint appliances.

Data Center 2 consists of two MDS 9222is running 3.2(2c) and a Switch Services Module running 3.2(3r) for SANTap services interfacing with the RecoverPoint appliances. The hosts and storage are connected to the integrated ports.

Figure 11 shows how RecoverPoint integrates into an existing storage fabric. On the left, there is a standard storage fabric setup. The server is connected to a storage switch as is a storage array. The SSM is installed in a MDS in the data path to the storage. With RecoverPoint added in, the physical connectivity of the server and host stays the same. The RecoverPoint appliances connect in to the MDS. Two vSANs are created instead of just one (for both host and storage). One zone is for frontend host zone and a Data Virtual Target (DVT), the backend zone contains the storage, Control Virtual Target (CVT), Appliance Virtual Target (AVT), and virtual initiators. These terms are briefly described below and in great detail in the EMC document referenced. Note that the hosts sit in a frontend vSAN and the storage and requisite RecoverPoint virtual instances reside in the backend vSAN.

Figure 11 Unidirectional RecoverPoint CRR Solution with SANTap



For more information, refer to EMC for *Deploying RecoverPoint with SANTap Doc 300-004-387 Rev A07*.

**Note**

When using RecoverPoint solution with SANTap, two vSANs are required for the solution. A frontend vSAN used for servers and a backend vSAN used for the RecoverPoint Appliances (RPA) and the storage arrays, which is shown by different colors for the different vSANs. Also virtual objects are created, such as DVT, CVT, and AVT.

The following subsection describes what each object is used for.

RecoverPoint-SANTap Terminology

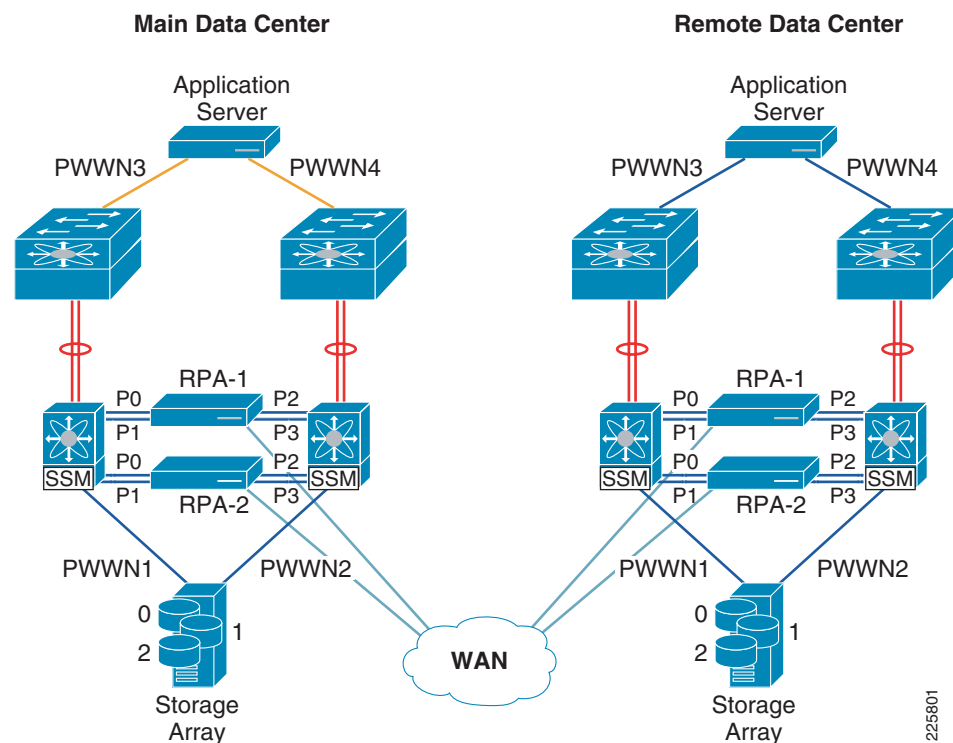
- *Backend vSAN*—This RecoverPoint Storage vSAN is considered the backend vSAN for all storage that will be used for RecoverPoint and will also be where the RecoverPoint Appliance Fibre Channel connections will reside when creating a CVI for SANTap, the MDS SSM will populate 9 virtual initiators plus the CVT into this backend vSAN. CVT is defined below.
- *Frontend vSAN*—Current implementation of SANTap requires that the host and storage ports that are using RecoverPoint with SANTap must be placed in two separate vSANs. Host initiators or host HBAs connections will be placed in the frontend vSAN.
- *Control Virtual Target (CVT)*—Each SSM that is enabled for SANTap can have multiple CVTs that is populated into the backend vSAN. This single CVT is used by the RecoverPoint appliances to communicate control functions from the appliances to the SSM for SANTap. The CVT is also synonymous with the RecoverPoint splitter. In the case where multiple backend vSANs are created. There is one CVT per backend vSAN.
- *Data Virtual Target (DVT)*—The DVT is a virtual target that resides in the frontend vSAN and mimics the physical storage port in the backend vSAN that is being used for RecoverPoint. This DVT allows the host to access the storage port as if the physical storage port resides in the frontend vSAN. For every storage port that is used for RecoverPoint, a DVT needs to be manually created by the SAN administrator. When a host in the frontend vSAN is zoned to the DVT, a host virtual initiator will be populated into the backend vSAN mimicking that host. A total of 16 DVTs can be supported with a single storage service module (SSM). There is a limitation in the number of hosts that can access the same DVT
- *Initiator-Target-LUN (ITL)*—This is basically the combination of a host HBA port, storage target port, and the LUN (disk) access through that storage port. Each unique ITL identifier will consume an ITL count on the SSM. This is relevant in regards that the number of ITLs supported per SSM is limited.
- *Appliance Virtual Target (AVT)*—The AVT allows to proxy communication to the LUNs that are being used for RecoverPoint. While the physical host is accessing the DVT in the frontend vSAN, the AVT creates a virtual initiators that is invisible to the end user, to obtain information about the LUNs. These AVT initiators are different from the host initiators and the 9 virtual initiators (described next) at the backend vSAN.
- *Virtual Initiator (VI)*—Enabling SANTAP feature on the SSM creates 9 virtual initiators that are populated into the backend vSAN. The 9 virtual initiators represent the 8 ASICs or Data Path Processors (DPP) on the SSM and the 9th virtual initiator represents the control processor on the SSM
- *Session*—A session consists of one of the ITL that is being used for replication or data protection. A host may have 10 ITLs, but maybe a subset of those ITLs are being used for data protection. Therefore, each ITL that is being replicated will consume a session.

Physical Requirements

EMC RecoverPoint best practice deploys clustered appliances (up to 8) per site. Within each RecoverPoint Appliance (RPA), there are two dual-ports FC HBA. Both ports on each of the HBA are connected to a single fabric. The Cisco MDS 9000 hardware requires one SSM per-physical fabric, per-site to be able to run SANTap. This means a minimum of two SSMs is required for the EMC RecoverPoint SANTap solution per-site. A total of four SSMs are needed, where each site has two SSMs for a complete RecoverPoint SANTAP CRR solution. There are no physical requirements to have RPAs hosts or storage ports directly connected to the SSM module to be able to use SANTap. EMC supports single fabric RecoverPoint solutions but as a best practice, dual fabric is recommended.

Figure 12 shows how Continuous Remote Replication looks with redundant storage fabrics at each site and redundant RecoverPoint Appliances. In this test scenario, the application is Exchange 2007 running on a VMware ESX server. Each server is redundantly connected to an MDS switch. Storage, an DMX 1000 in this setup, is homed to two MDS switches. The RecoverPoint appliances connect to both MDS switches in data centers 1 and 2. The WAN ties into the access layer and is described in the switch/router setup. Figure 12 shows that, with an existing WAN and datacenter infrastructure, RecoverPoint can be easily added in without an entire redesign and moving physical ports.

Figure 12 *Unidirectional RecoverPoint CRR Solution with SANTap*



SAN Design Considerations for EMC RecoverPoint

There are a few network design considerations that need to be taken into account when deploying EMC RecoverPoint with SANTap. Deployments will vary depending on the physical topology of each environment. The following things need to be considered:

- Where should the SSM be placed in the network?
- Should anything be connected to the SSM front-panel ports?

- Where should the RecoverPoint Appliance Fibre channel ports connect?
- Will SANTap work with IVR?

Placement of the SSM for SANTap

As a best practice, placing the SSM in the switch nearest to where the storage devices are located is recommended. In most large SAN topologies, the storage devices will be servicing multiple servers, where those servers may span across different switches. This core-edge topology places the storage as a service to the servers. Placing the SSM nearest to the storage devices, which is at the core, allows the SANTap feature to service more servers easier without adding unnecessary network traffic. Within a core-edge topology where both servers and storage are connected at the edge and the core is used for SAN traffic to allow communication to other edge switches, it is recommended that the SSM should be placed at the edge switch for the storage devices that are using SANTap. In topologies like this, the edge switches are used for localized traffic and will have minimal traffic traversing to the core.

SSM Front-Panel Ports

With the SSM allowing for advanced features such as SANTap and also having the capability to provide Layer 2 FibreChannel switching, this becomes a factor in designing a SANTap solution. Each of the 8 ASICs that use SANTap shares 2.5Gbps of bandwidth with the corresponding 4 front-panel ports. With this in mind, it is clear that better throughput performances are achieved if there are no devices connected to the front-panel ports of the SSM.

When connecting devices to the SSM front-panel ports it is important to understand the recommendations. Every 4 ports on the SSM and the associated ASIC, which is called the Data Path Processor (DPP), share 2.5Gbps of bandwidth. It is recommended not to plug storage devices into these ports because storage devices typically are servicing many servers and will fully use 2Gbps of bandwidth per port. If storage devices will be connected, it is recommended to plug only one storage device on every 4th port. Therefore, use ports 1, 5, 9, 13, etc. for storage devices and leave the other front-panel ports unused. Whereas if you plug servers or RPAs on any of the front-panel ports, most servers typically do not require that much bandwidth and would not be as affected as storage ports in this type of solution design. If all ports in the 4-port group are connected by servers, those servers and any other servers using SANTap accessing that DPP will be sharing that 2.5Gbps of bandwidth. There are environments that have storage, servers and RecoverPoint appliances all connected to the ports on the SSM. This is not recommended but is possible to implement. For support of this type of implementation, consult EMC RecoverPoint engineering team.

RecoverPoint Appliance Connectivity

As a best practice, connecting the RecoverPoint Appliance FC ports on non-oversubscribed ports is recommended. Since the appliance will be servicing multiple servers and storage ports, it makes most sense to have the appliance ports connected near the storage ports as well as on full-line rate ports that can handle 2Gbps of bandwidth. In a core-edge topology where servers are connected at the edge and storage connected at the core, treating the appliance connection as storage is recommended.

SANTap with IVR Considerations

Despite having servers in the frontend vSAN and disks in the backend vSAN, there is no need of IVR feature because SANTap will make sure there is communication across the two vSANs through virtual entities. IVR can be used for non-SANTap hosts. For example, a SANTap host initiator is accessing its storage device in the backend vSAN but also needs access to a tape device, which is also located in the backend vSAN. Since SANTap does not work with tape devices, one could be tempted to use IVR to

resolve this issue. However, since SANTap creates in the backend vSAN, a virtual initiator that mimics the real host initiator, IVR will not work because it would create a communication loop. There are two workarounds for this situation:

1. Move the tape device to the frontend vSAN so that no IVR is needed.
2. Create another vSAN for the tape device, different than the backend vSAN, and then use IVR for the host initiator located in the frontend vSAN to route to the tape device.


RecoverPoint Consistency Group

RecoverPoint should be installed, configured, splitters attached, storage and hosts zoned properly, and arrays masked correctly so that the appliances can see the LUNs available in each data center for replication.

-
- Step 1** Create a consistency group by selecting **Groups** from the main menu bar of the RecoverPoint Java Client. It will present what is shown in [Figure 13](#). Since ESX was used in this setup, reservation support is necessary. Errors will occur and the **Consistency Group** will not function properly if this is not selected. Also note the **Compression** tab. RecoverPoint offers 10 levels of compression to balance compression and performance. RecoverPoint offers 10 levels of compression to balance compression and performance. This test setup did not use compression on the RecoverPoint appliances. Instead, the CPU cycles on RecoverPoint are left to replication, and Cisco WAAS is used to compress the replication data on the WAN. All of the values shown in [Figure 13](#) are accessible for modification or review after creation.

Figure 13 RecoverPoint Consistency Group Configuration

Create a Consistency Group
Define the properties of the new consistency group.



General Settings

Name :

Preferred RPA:

☒ Reservations Support

Compression

Replication data can be compressed at various levels prior to transfer over the WAN.

☐ Enable Compression Compression Level : N/A

Best Compression Best Performance

Optimization

RecoverPoint can utilize as much bandwidth as necessary to keep lag to a minimum, or expend bandwidth only when needed to keep lag within its allowed range. It can also slow or halt application(s) upon nearing a policy bound.

Minimize : ☒ Lag ☐ Bandwidth

Resource Allocation

Bandwidth can be adjusted, and a priority vis-a-vis other consistency groups can be set.

Priority :

Bandwidth Limitation : ☒ Unlimited ☐ Limited Mbps

Advanced

General

Hosts OS :

Global cluster mode :

☐ Allow Regulation

☒ Allow data transfer even when group is handled by non-preferred RPA

☒ Measure lag when writes reach the target RPA (as opposed to the journal)

CDP Specifications

☒ System Optimized Lag

☐ Lag : MB

Snapshot Granularity :

☒ Perform fast first-time initialization

CRR Specifications

☒ System Optimized Lag

☐ Lag : MB

Snapshot Granularity :

☒ Perform fast first-time initialization

225602

- Step 2** Production copy settings are configured on the next screen. As shown in [Figure 14](#), the defaults were accepted: no journal compression, and no specific protection window. Depending on size of journal space available and recovery time objective (RTO) /recover point objectives (RPO), these two options may need to be changed to fit the recovery requirements for an application.

Figure 14 **Configure Production Copy Settings**

New Consistency Group Wizard

Production Copy Settings
Define the production copy of this group.

Select Site: sj1

General Settings

Name : Exchange 2007

Journal Compression Level : None

☐ Required protection window: 1 hours

Advanced

Maximum Journal Lag:
☒ Unlimited
☐ Limited
 0 bytes

Proportion of journal allocated for target-side log (%):
 20

Journal size limit (GB):
 1200

Reservations Policy:
 Auto

☒ Allow distribution of snapshots that are larger than capacity of journal volumes.

< Back Next > Finish Cancel

2259003

Step 3 The next screen, as show in [Figure 15](#), configures the replication settings. Do not configure local copy settings. This is Continuous Data Replication (CDP), which creates a local copy on the existing array and allows significant flexibility when it comes time to recover or roll back to different images. Currently, CDP does not work with VMware Site Recovery Manager and the Storage Recovery Adapter for RecoverPoint and must not be selected.

Figure 15 *Configure Replication Settings*

Step 4 Replication sets are configured for each volume that needs to be replicated. In this Exchange setup the ESX server has three LUNs. The main OS Storage LUN is 45Gb, the message store 140Gb, and the log 80Gb. This requires three replication sets within the one consistency group. The LUNs are combined in one group to keep all three LUN copies synchronized with each other. It would be of no use to have a log that is 10 minutes older than the current message store. Note that if additional storage is added to the Exchange VM in the future, it can be added to the existing consistency group as a new replication set.

Volumes are added with the add **volume** tab. RecoverPoint will present the available LUNs it can see. If masking is not done properly or is not complete, volumes will not appear. The setup will not allow the configuration to continue if the LUNs are not added to the replication sets or the journals. For each set, a local LUN is selected (the production source) and then select the recovery side copy of this LUN must also be selected for the CRR volume. The remote LUN can be larger than the replicated LUN, but not smaller. Any extra space will remain unused. The test setup recovery LUNs were larger at the recovery site and the protected site and no problems were encountered with replication or recovery.

Add volumes to the local journal and the CRR journal on the remote recovery side. The local journal is used for failback. Journal size calculations can be calculated with a formula. Refer to the EMC RecoverPoint Administrator's Guide (available on PowerLink) for more information on calculating journal size.

The recommended journal size equals as follows:

$$1.05 * [(? \text{ data per second}) * (\text{required rollback time in seconds}) / (1 - \text{target-side log size as fraction})]$$

If the following values are used:

- ? data per second = 5 Mb/s
- required rollback time = 24 hr = 86 400 s
- target-side log = 0.20

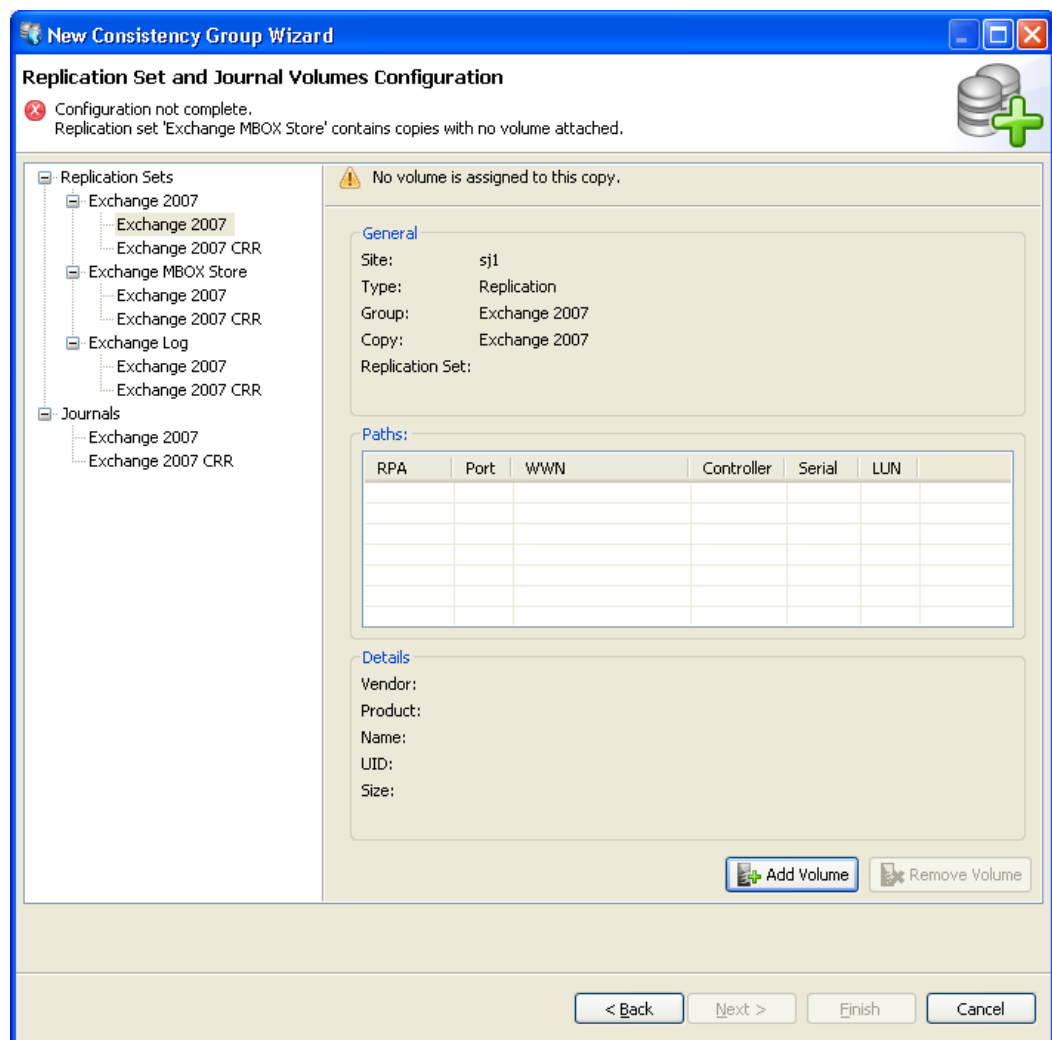
Then the required journal size equals as follows:

$$1.05 * 5 \text{ Mb/s} * 86\,400 \text{ s} / (1 - 0.20) = 567\,000 \text{ Mb}$$

$$567\,000 \text{ Mb} = 567\,000 / 8 \text{ MB} = 70\,875 \text{ MB} = 70.9 \text{ Gb}$$

Step 5 When finished, **RecoverPoint** configures an instance of the consistency group. An example of the configuration is show in [Figure 16](#).

Figure 16 *Configure Replication Set/Journals and Attach Storage Volumes*

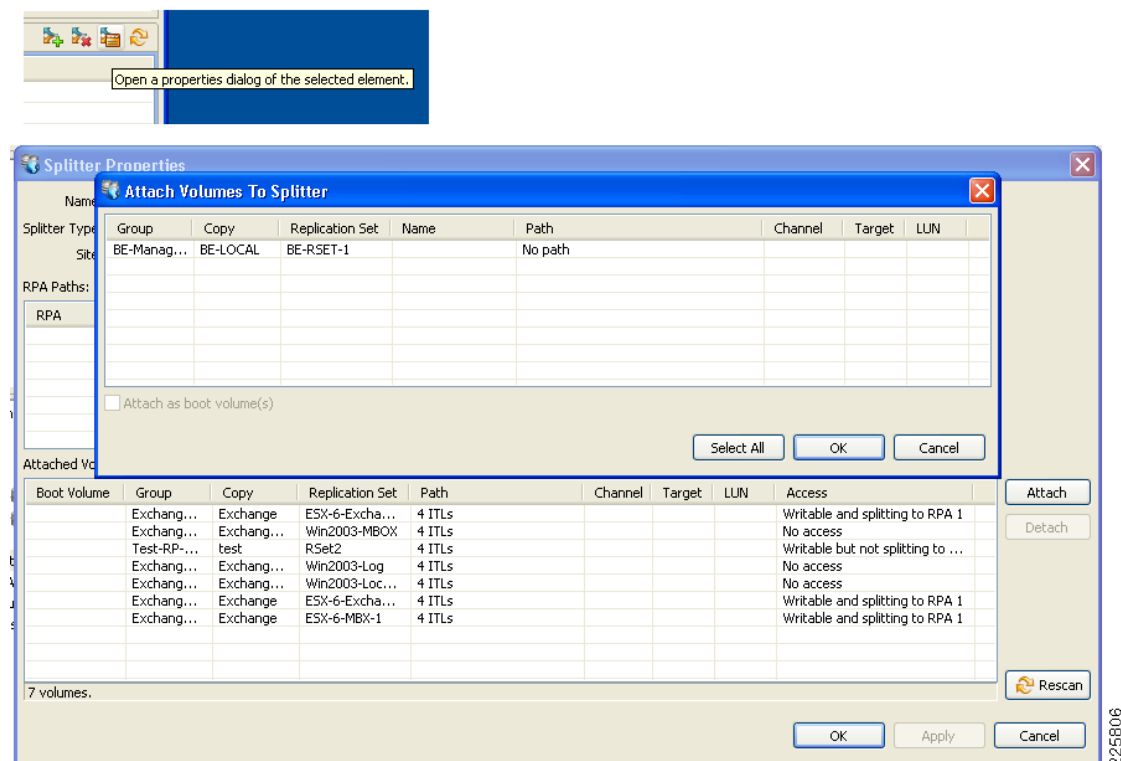


2259/05

Step 6 On the side menu, select splitters and select the business card just above the splitter box on the right. Press the **attach** button and select your unattached volumes that were just assigned to the replication sets. Attach the volumes and repeat for each splitter. Splitters can be referenced during troubleshooting to verify SANTap and RecoverPoint are working.

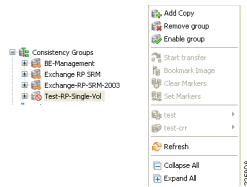
See [Figure 17](#) for an example of attaching volumes to a splitter. The splitters in this setup are in the Storage Services Modules inside the MDSSs. These modules are running SANTap services. For the traffic to go from a host in the frontend vSAN to storage in the backend vSAN it traverses the splitter. Select a splitter for each Replication.

Figure 17 *Attach Volumes to Splitters*



Step 7 Consistency groups must be activated after configuration. RecoverPoint will put the group into a full sweep initialization where it compares by block the production source to the recovery LUN. Initialization can take a long time depending on how fast the WAN is and how large the replication LUNs are. In the Exchange Server setup, there are 265Gb of data to be synchronized between the two storage arrays. Synchronizing over 3MB or two T1s will take significantly longer than over a Gb link. Even on a Gb link it can take several hours. Enable consistency group as seen in [Figure 18](#). If the group is not enabled, synchronization will not start, nor will replication.

Figure 18 **Enable Consistency Group**



- Step 8** Recover Point is now setup for Site Recovery Manager. The only other setup that requires RecoverPoint credentials is setting up the Storage Recovery Adapter Array Manager in Site Recovery Manager. The RecoverPoint GUI can be referenced for status and utilization during testing but no interaction is needed because SRM manages all the image recovery through the Storage Recovery Adapter that directly interfaces the RecoverPoint appliances.

Cisco Design and Deployment

Site Selection

Site selection (or content routing) provides user recovery services associating a single user with an application instance. In [Figure 10](#), the Cisco GSS provides this DNS-based service for the Exchange environment. The GSS appliances are the external DNS authoritative Name Servers for the enterprise providing A records for the Cisco ACE VIPs associated with the CAS role. The GSS appliances are deployed at the edge of the network as a cluster. Clustering the GSS boxes across multiple enterprise sites provides for a highly available and scalable DNS solution. It is important to note that each GSS can house an instance of Cisco's Network Registrar (CNR) that supplies the Mail Exchanger (MX) Resource Records to properly route Internet mail into the enterprise. In this document, the GSS is using an existing DNS infrastructure for MX records.

Typically, enterprise deployments use DNS-based round-robin and multiple MX records to distribute load across multiple Internet mail servers. This method provides redundancy and scalability but relies heavily on the client mail application to resolve DNS and MX records for mail delivery. The sample output below shows two SMTP gateways that have equal MX preferences and therefore, both will be used in a DNS-based round-robin fashion.

```
ese.com MX preference = 10, mail exchanger = ironport1.ese.com
ese.com MX preference = 10, mail exchanger = ironport2.ese.com
ironport1.ese.com      internet address = 10.5.25.11
ironport2.ese.com      internet address = 10.6.25.11
```


To provide a higher level of scalability and availability for inbound and outbound mail processing, administrators may choose to load-balance across multiple IronPort appliances. A load-balancer, such as the Cisco ACE, provides high availability and scalability within a single data center and is able to communicate the state of the IronPort appliances to the Cisco GSS that are globally aware. The GSS may probe the ACE VIP at each site using multiple probe types. Combining these probes allows the GSS to gain better insight into the state of the services located at each data center, providing a more intelligent form of SMTP server (i.e., site selection).

Traffic Pattern

[Figure 19](#) illustrates the DNS traffic patterns and probing technology of the GSS. In this example, the GSS is actively probing the Cisco ACE service modules located at two distinct data centers in the enterprise. The Cisco GSS is aware of the IronPort application state. It reflects this knowledge in its resolution of DNS requests.

In this example, the GSS probes the ACE virtual context VIP hosting the ET roles using KAL-AP. As defined earlier, KAL-AP provides load and VIP state information for each VIP maintained within a context hosting a KAL-AP agent. To enable KAL-AP agent on the ACE, use the following commands:

```
#Used for return connection to GSS (key must match CAPP Hash Secret on GSS)
kalap udp
    ip address 10.5.25.10 encryption md5 ese
!
#Permit GSS for KALAP-UDP connection (optionally this can permit "any" instead of specific
#GSS(s))
class-map type management match-any KALAP-MGMT
    2 match protocol kalap-udp source-address 10.5.25.10 255.255.255.255
policy-map type management first-match KALAP-MGMT-POL
    class KALAP-MGMT
        permit
!
#Apply policy to incoming interface
interface vlan 103
    description PUBLIC-FACING
    service-policy input KALAP-MGMT-POL
!
#IP address on BVI used by GSS for Shared VIP KeepAlive
interface bvi 1
    ip address 10.5.103.4 255.255.255.0
```

The GSS configuration requires one to define a shared keepalive of type KAL-AP and associate it with the KAL-AP agent defined on the ACE via the primary or secondary IP address fields, [Figure 19](#) shows an example of this relationship. The 10.5.103.4 is an IP address defined on the bridged virtual interface of the virtual ACE context.

Figure 19 Example GSS Configuration of Shared KAL-AP

Modifying Shared VIP KeepAlive: 10.5.103.4	
Type:	<input type="radio"/> ICMP <input type="radio"/> TCP <input type="radio"/> HTTP HEAD <input checked="" type="radio"/> KAL-AP <input type="radio"/> SCRIPTED KAL
Shared KAL-AP KeepAlive	
Primary IP Address: *	10.5.103.4
Secondary IP Address:	
CAPP Secure:	<input checked="" type="checkbox"/>
CAPP Hash Secret:	ese
Fast KeepAlive Settings	
Number of Retries:	5 Range: 1 - 10
Number of Successful Probes:	1 Range: 1 - 5

KAL-AP VIP configuration screen allows one to set variables specific to the KAL-AP probe or one may leverage global parameters, timers and security settings. It is recommended to secure all communication between the GSS and ACE devices, to force this practice use the CAPP Hash Secret globally on the GSS.

**Note**

The CAPP Hash Secret is equivalent to the *kalap udp* agent password on the ACE configuration.

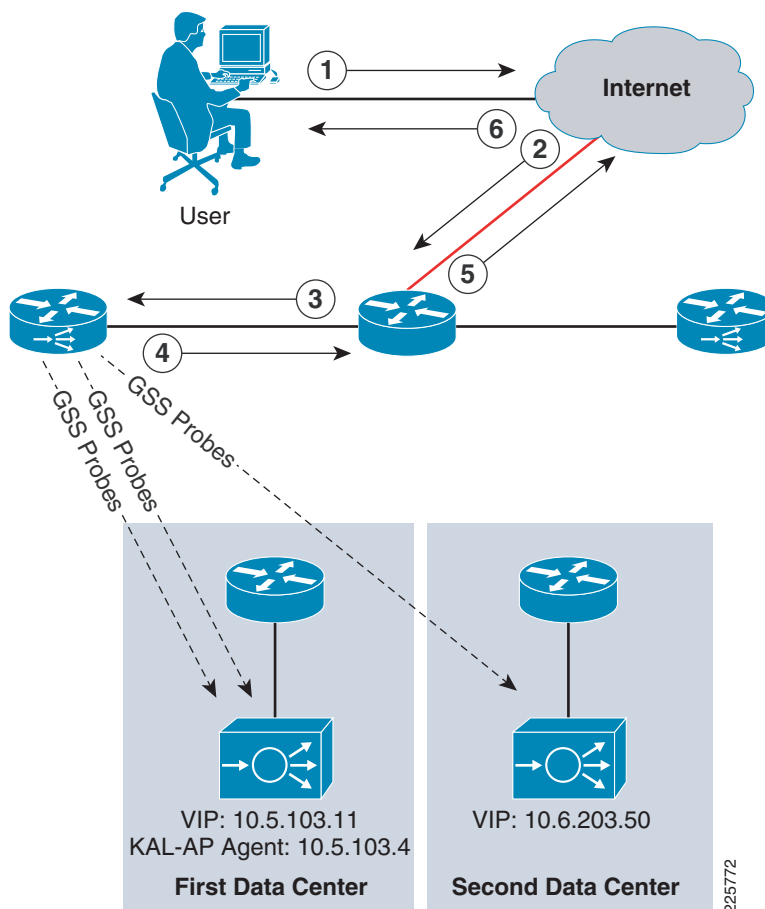
To allow the GSS to apply the KAL-AP knowledge to an associated A record representing the ACE VIP, assign the KAL-AP keepalive to the Answer (see [Figure 20](#)). Note that the VIP address and KAL-AP keepalive reference IP addresses defined on the ACE context.

Figure 20 Example GSS Keepalive to Answer Configuration

Modifying Answer: DC1-OWA-VIP / 10.5.103.11	
Type:	VIP
Name:	DC1-OWA-VIP
Location:	Data Center 1
Manual Reactivation:	<input type="checkbox"/>
VIP Answer	
VIP Address: *	10.5.103.11
VIP KeepAlive Type:	<input type="radio"/> None <input type="radio"/> ICMP <input type="radio"/> TCP <input type="radio"/> HTTP HEAD <input checked="" type="radio"/> KAL-AP <input type="radio"/> SCRIPTED KAL <input type="checkbox"/> Multi-port
KAL-AP KeepAlive	
KAL-AP Type:	KAL-AP By Vip
Shared KAL-AP KeepAlive:	10.5.103.4
Tag:	

With the above GSS and ACE configurations completed, basic traffic flow can be tested. [Figure 21](#) illustrates the DNS traffic patterns and probing technology of the GSS. In this example, the GSS is actively probing the Cisco ACE service modules located at two distinct data centers in the enterprise.

Figure 21 GSS OWA Request and Probe Flow



The following steps outline the external mail server requests for MX records:

- Step 1** An external OWA user issues a standard DNS query for owa.esec.com.
- Step 2** The router receives the UDP packet and forwards to the GSS the authoritative name server for the domain.
- Step 3** The GSS receives the DNS request.
- Step 4** The GSS responds with the resource record for owa.esec.com. This resource record configuration directs the incoming traffic to the owa.esec.com server. The owa.esec.com server resolves in DNS to the ACE VIP.
- Step 5** The router forwards the packet.
- Step 6** The external OWA user receives the record from the Cisco GSS, the authoritative name server for the domain.

There are many options for performing site selection with the Cisco GSS. The GSS uses what is known as a *Balance Method* to determine its site selection order for a given policy. There are many options under the Balance Method configuration option such as **Hashed**, **Least Loaded**, **Ordered List**, **Round Robin**, and **Weighted Round Robin**. Within these balance methods, additional decision making is available by adding proximity capabilities.



Note

In the testing conducted for this design the **Ordered List** and **Least Loaded** options were validated.

The **Least Loaded** option enables the GSS to reply to “A” record queries for the VIP associated with owa.es.com based on the VIP load reported by the Cisco ACE KAL-AP agent. There is an Answer Group associated with Data Center 1 and the VIP associated with owa.es.com is in Data Center 1. The GSS will monitor the KALAP-UDP connections for availability/load of both data centers and if the VIP in Data Center 1 becomes too loaded (based on the configured **Load Threshold**) or unavailable, the GSS will then change its record information to use the Answer Group for Data Center 2. **Least Loaded** (also the other options such as **Round Robin**) will use an active/active deployment as the GSS will simply send traffic to either the least loaded sites in a alternate fashion or make a basic round robin (or weighted) decision.

Table 2 illustrates the Cisco GSS configuration for owa.es.com using **Least Loaded**.

Table 2 Cisco GSS Configuration Summary

Cisco GSS Configuration Option	Name/Type	Setting
Source Address List	Anywhere	0.0.0.0/0 (Any address)
Domain Lists	OWA-DL/Domains	owa.es.com
Answer Groups	OWA-AG/VIP	VIP=10.5.103.11, Name=DC1-OWA-VIP, Location=Data Center 1, Order=0, Load Threshold=254, Weight=1 VIP=10.6.203.50, Name=DC2-OWA-VIP, Location=Data Center 2, Order=1, Load Threshold=254, Weight=1
Answers	DC1-OWA-VIP/VIP DC2-OWA-VIP/VIP	Location=Data Center 1, KeepAlive Method=KAL-AP by VIP Location=Data Center 2, KeepAlive Method=ICMP to VIP
Shared KeepAlives	10.5.103.4/KAL-AP	CAPP Secure

Figure 22 shows the DNS Rule that ties all of the above together.

Figure 22 DNS Rule

Name	Source Address List	Domain List	Owner	Status	Answer Groups	Balance Methods	Clause Status
OWA-DR	Anywhere	OWA-DL	System	Active	1: OWA-AG	1: Least Loaded	1: Active

Showing 1-1 of 1 records

The **Ordered List** option is not a favorable choice if a true active/active data center design is the goal. This is true because **Ordered List** works within the boundaries of the Answer Groups’ *Location* and *Order* values. This means, in the configuration used in this testing, the **Ordered List** will always use the VIP **10.5.103.11** (Order 0) in Location **Data Center 1**. If the VIP is also loaded (beyond 254) or down, then the GSS will use the next entry in the Ordered List—VIP **10.6.203.50** (Order 1), Location **Data Center 2**.

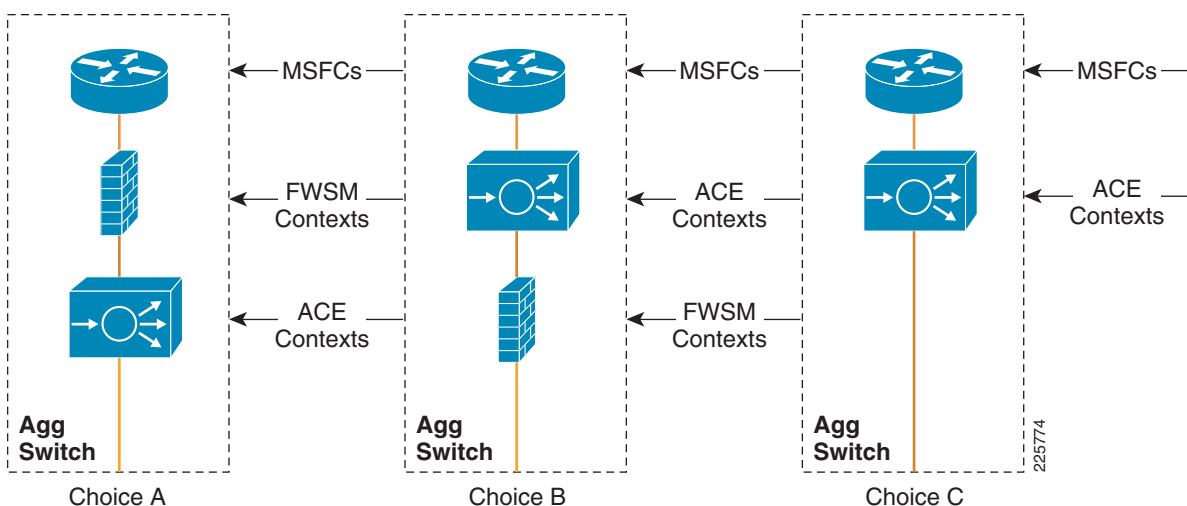
Cisco ACE Deployment for the CAS Role

This section focuses on the deployment of the CAS role in the data center. As previously mentioned, the CAS Exchange server roles is an excellent candidate to take advantage of network-based services such as load balancing, SSL offload, firewall and other application optimization services. Figure 23 below highlights the flexibility and some basic design options afforded to the enterprise leveraging Cisco’s virtual service technologies. These design options provide centralized load balancing, security, and optimization services for the application. In addition, the virtualization capabilities of both the FWSM

and the ACE allow a single physical device to provide multiple logical devices to support a variety of application environments. System administrators can assign a single virtual device to a business unit or application to achieve application performance goals, requirements, or service-level agreements.

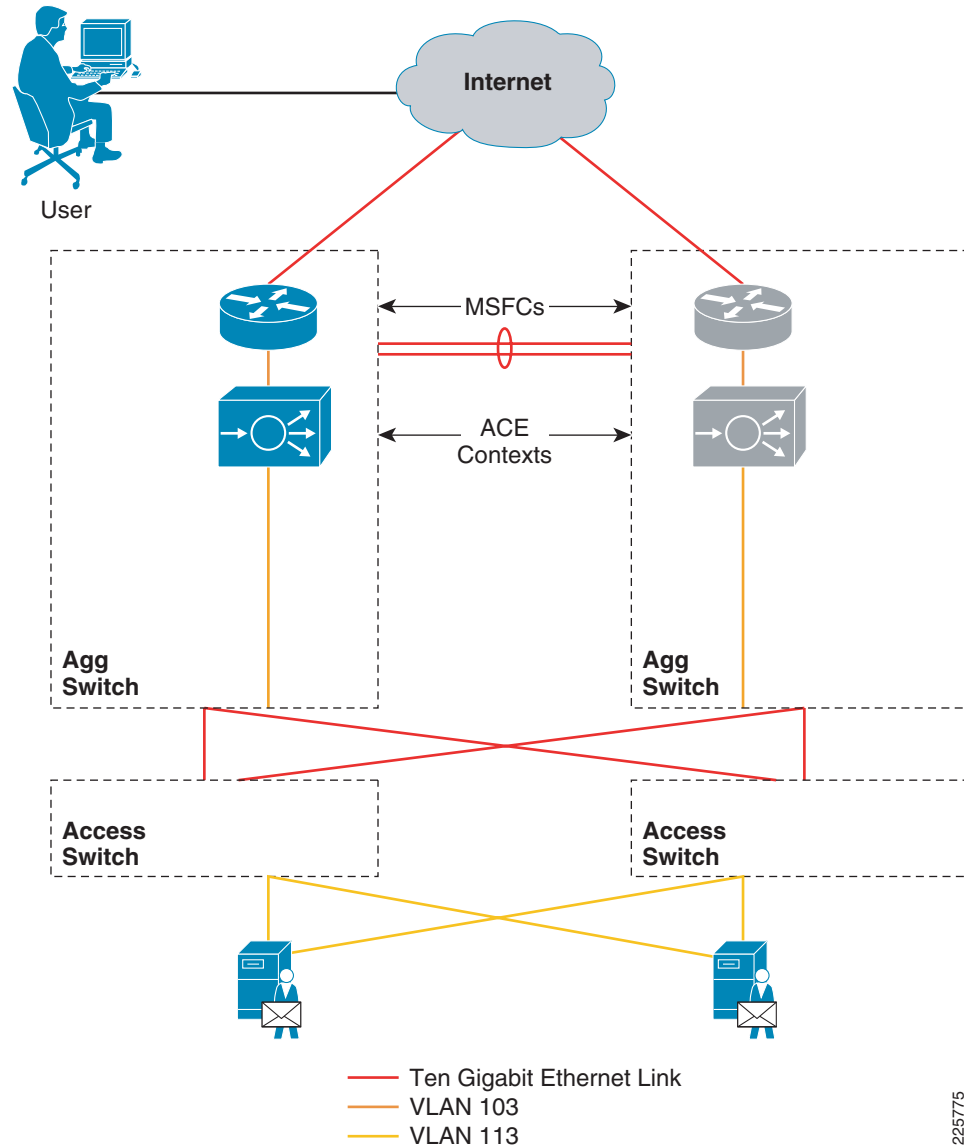
In [Figure 23](#), the FWSM and ACE virtual contexts are logically positioned north and south of one another. Choice 'A' places the FWSM virtual context as a traditional firewall controlling access to the CAS and network services cascading to the south. Choice 'B' uses the firewall security services available in the Cisco ACE to provide protection and other application services. The FWSM context below the ACE context is another layer of protection into the internal enterprise. The final option depicted, choice 'C', relies on the security and application services of the ACE virtual context to protect and optimize the data center applications.

Figure 23 Enterprise Data Center Design Choices – ACE and FWSM



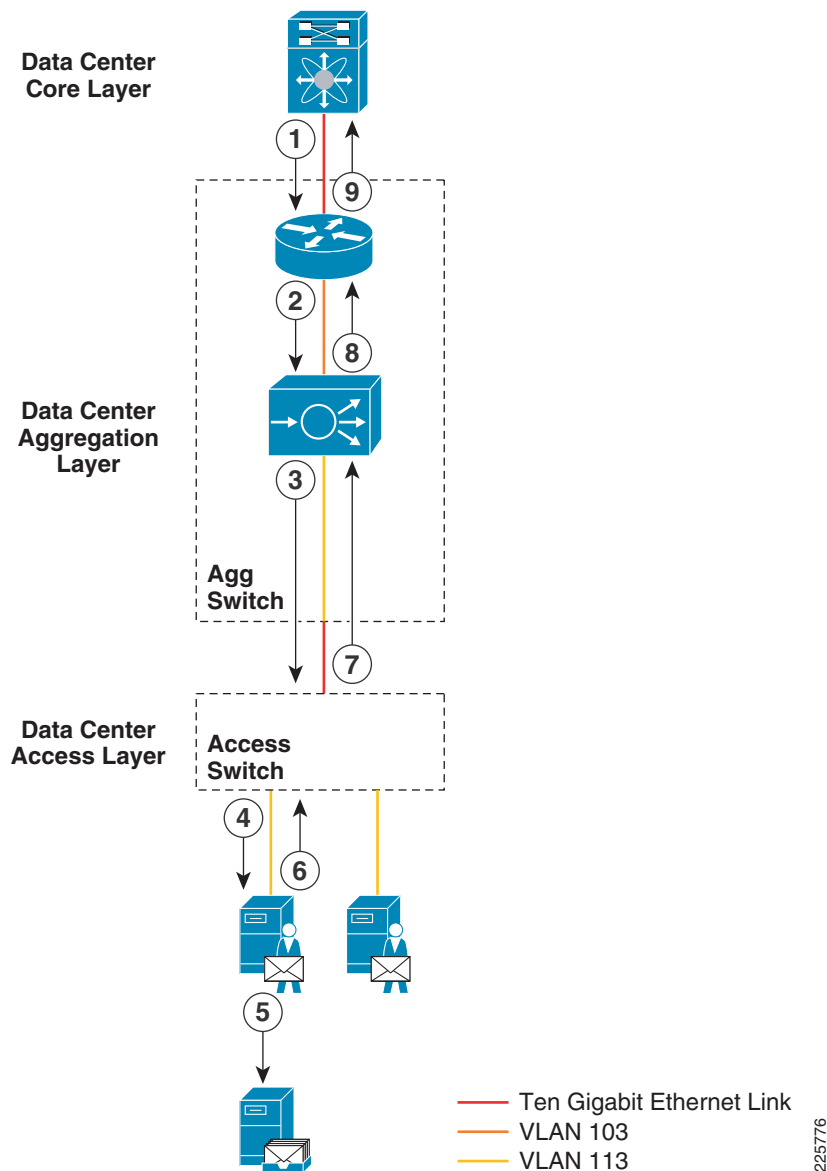
In [Figure 24](#), each Catalyst 6500 aggregation switch is an ingress point to the enterprise data center access layer providing transport, application, and security services. The ACE virtual contexts residing in the Catalyst platform support firewall functionality, load balancing, and applications acceleration. The Microsoft Exchange 2007 CAS role is positioned behind a dedicated ACE virtual context explicitly configured to support each role. [Figure 24](#) shows the public, or client facing VLAN as VLAN 103 and the internal or server facing VLAN as VLAN 113. The ACE creates a bridged virtual interface, BVI, to associate the two VLANs to one another. The CAS roles point to the MSFC as their default gateway.

Figure 24 Cisco ACE VLAN Layout – CAS Role



Traffic Pattern

Figure 25 depicts the flow of OWA or Outlook Anywhere traffic in and out of the data center.

Figure 25 *Transparent Service Integration with CAS*

The integration of Microsoft Exchange's CAS roles and the ACE in transparent mode includes the following steps:

-
- Step 1** The data center core routes incoming traffic to the ACE VIP.
- Step 2** The MSFC of the aggregation layer receives the traffic and forwards it to the ACE VIP defined on VLAN 103 in this example. The ACE is in transparent mode and therefore leverages a bridged virtual interface to join VLANs 103 and 113. The following interface configuration is necessary:

```
interface vlan 103
description OWA-Client-side-vlan
#Joins the interface to the bridge group
bridge-group 1
#This service policy defines the CAS VIP and the ACE contextual services it offers.
service-policy input OWA-POLICY-MAP
```

```

no shutdown

interface vlan 113
description Server-side-vlan
#Joins the interface to the bridge group
bridge-group 1
no shutdown

#This is the logical join between the two VLANs
interface bvi 1
ip address 10.5.103.4 255.255.255.0
description OWA-Bridged-vlans
no shutdown

```

- Step 3** The ACE bridges the traffic to the CAS roles defined in its server farm after applying the associated load balancing, and security policies. The destination IP address becomes the selected CAS roles IP address, but the source address remains that of the originating client.
- Step 4** The CAS role receives the flow with the clients IP address.
- Step 5** The CAS role communicates directly with the appropriate mailbox server for the user.
- Step 6** The CAS sends flow back to its default gateway the MSFC.
- Step 7** The ACE receives the packet on the VLAN 113 interface transparently bridging the communication from the CAS to the MSFC.
- Step 8** The ACE forwards the flow to the MSFC, changing the source IP address from the CAS role to its VIP. The client is unaware of the “real” CAS role IP address.
- Step 9** The MSFC routes the flow back to the client.
-

ACE for OWA

The CAS role supports browser based user connectivity to the Microsoft Exchange environment via a web browser. This thin client implementation has become increasingly popular and more functional for remote users. The CAS role services are available via Microsoft’s Internet Information Server (IIS) as a virtual directory named “owa”, therefore port 80 and 443 are the default points of access to the environment. In the tested configuration, the ACE module frontends the environment providing the following services:

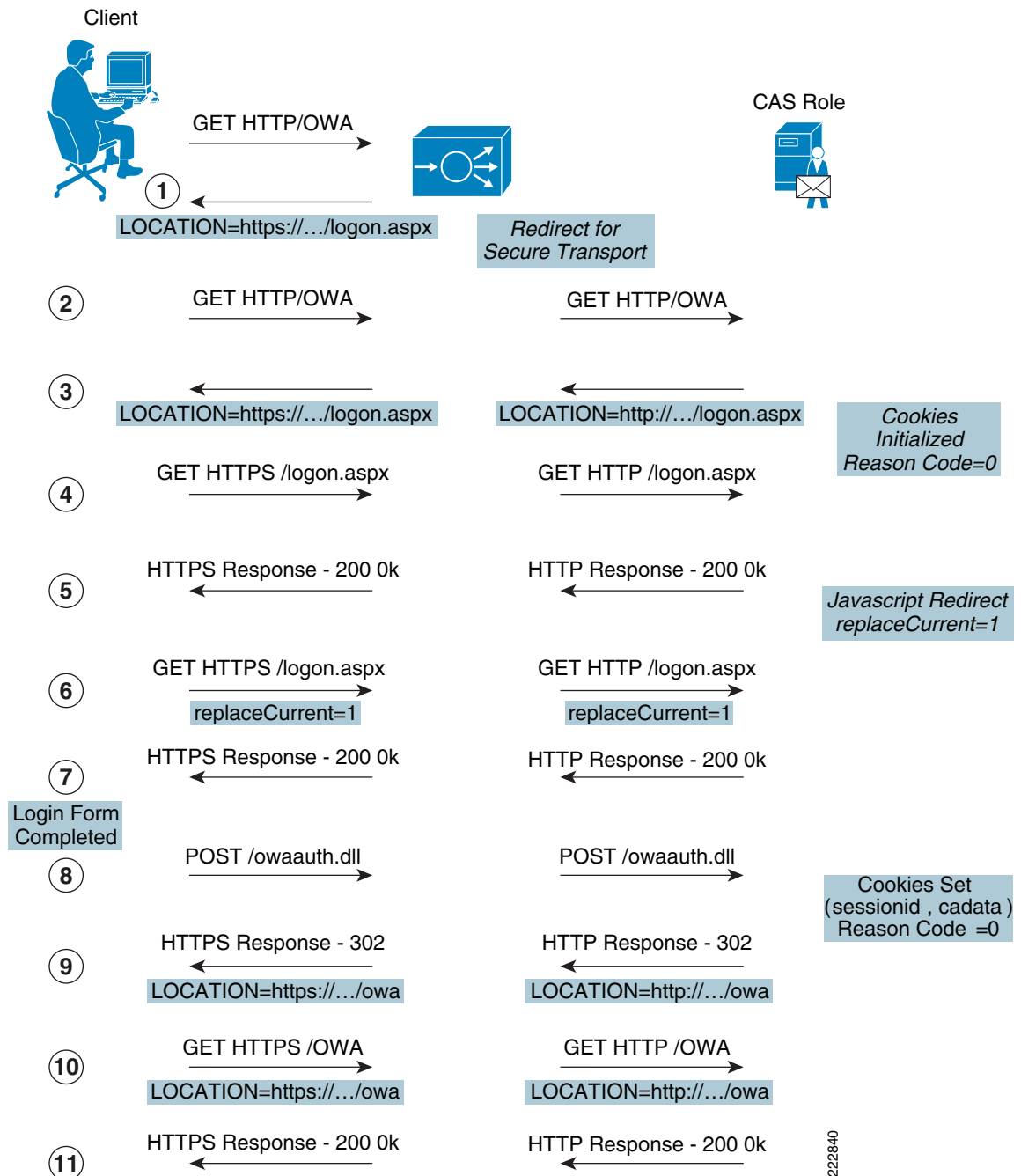
- SSL-offload
- Access control
- Load balancing
- Session Persistence
- Health Monitoring
- Server Redirect
- URL Rewrite

To better understand the services provided via the ACE module in an OWA environment, it is best to look at an example. [Figure 26](#) depicts an OWA logon session and will be used to describe the use of ACE services for OWA.



In this example, the client attempts to connect via HTTP but it is redirected to HTTPS as defined by the security policies of the enterprise. In addition, the Cisco ACE is positioned as an SSL proxy, offloading the SSL processing from the CAS role.

Figure 26 **Example CAS OWA Logon**



In [Figure 26](#), the following flow occurs:

- Step 1** The client browser attempts to access the OWA application at <http://ese.com/owa> in DNS this resolves to the ACE VIP defined in a class-map for ports 80 (HTTP) and 443 (HTTPS).

```
policy-map multi-match OWA-POLICY-MAP
  class OWA-VIP
    #Enable the virtual IP address
    loadbalance vip inservice
    #Redirect load balance policy from port 80 to 443
    loadbalance policy REDIRECT-PM
    #Allow VIP to respond to pings
    loadbalance vip icmp-reply active
    #Enable RHI
    loadbalance vip advertise active
    #Enable TCP reuse feature of ACE
    appl-parameter http advanced-options TCP-REUSE

  class OWA-VIP-443
    #Enable the virtual IP address
    loadbalance vip inservice
    #Session persistence and load balance policy
    loadbalance policy STICKY_IP_LB
    #Allow VIP to respond to pings
    loadbalance vip icmp-reply active
    #Enable RHI
    loadbalance vip advertise active
    #Enable TCP reuse feature of ACE
    appl-parameter http advanced-options TCP-REUSE
    #Enable ACE to perform SSL Services
    ssl-proxy server SSL-OFFLOAD
```

The client-side interface, in this case VLAN 103, uses the OWA-POLICY-MAP policy map:

```
interface vlan 103
  description OWA-Client-side-vlan
  bridge-group 1
  no normalization
  #Apply policy map to interface
  service-policy input OWA-POLICY-MAP
  no shutdown
```

The policy map OWA-POLICY-MAP has a load-balance policy name REDIRECT-PM under the class OWA-VIP. This class defines the VIP listening under port 80. The system administrator, in this example, wants all incoming traffic to leverage secure transport services, i.e. SSL. The ACE redirects the client via an HTTP 302 message to use the following LOCATION as defined in the HTTP header, <https://owa.ese.com/owa>. The CAS_REDIRECT policy uses a class of default allowing it to catch all traffic destined to the VIP on port 80.

```
#Catch all port 80 traffic to the CAS-VIP and send to OWA redirect serverfarm
class class-default
  serverfarm REDIRECT-SERVERFARM

#A single redirect server is defined
serverfarm redirect REDIRECT-SERVERFARM
  rserver REDIRECT-TO-HTTPS
  inservice

#The redirect server defines the correct URL to access the OWA services of Exchange
#via a 302 HTTP header location statement
```

```
rserver redirect REDIRECT-TO-HTTPS
  webhost-redirection https://owa.esa.com/owa 302
  inservice
```



Note The redirect server may also use HTTP 301 messages. HTTP 301 messages are typically given in response to client POSTs and may be rejected by some browsers if received after a client GET message.

- Step 2** The client issues a secure HTTP GET for the new location at <https://owa.esa.com/owa>. This ACE policy map OWA-POLICY-MAP supports this VIP defined under the OWA-VIP-443 class map. In addition, the ACE provides SSL-offload for the CAS roles with the SSL proxy server configuration. The ACE negotiates a secure session between itself and the client. The following commands define the SSL proxy service:

```
ssl-proxy service SSL-OFFLOAD
#Identifies the RSA key used for the OWA certificate
  key testkey.key
#Defines the certificate issued by the trusted provider
  cert testcert.pem
```

In addition to the SSL proxy service, the ACE provides load balancing through a load-balance policy named STICKLB. The STICKLB policy provides both a load balancing and a session persistence mechanism. The default load-balancing algorithm is round-robin. The session persistence method, in this case, uses a cookie inserted by ACE named **ACE-Insert** against the OWA server farm named OWA. The configuration is as follows:

```
#Define the OWA server farm
serverfarm host OWA
#ICMP alive Probe
  probe ping
#Real CAS servers listening on port 80
rserver CAS1 80
  inservice
rserver CAS2 80
  inservice

#HTTP cookie persistence using the inserted cookie named "ACE-Insert"
sticky http-cookie ACE-Insert sticky-cookie-group
#Set Cookie inactivity timeout (in minutes)
  timeout 20
#CAS server available for service
serverfarm OWA
```



Note The default inactivity cookie timeout for OWA is 20 minutes. In this example, the ACE cookie inactivity timeout is also set to 20 minutes to remain consistent with the Microsoft CAS role deployment.

```
#This policy map defines the load balancing for all SSL traffic destined to the OWA farm
#with session persistence based on the ACE inserted cookie "ACE-Insert"
policy-map type loadbalance http first-match STICKYLB
  class class-default
    sticky-serverfarm sticky-cookie-group
```

The ACE sends clear text HTTP to one of the OWA servers defined in the serverfarm.

- Step 3** The CAS role receives the HTTP GET from the ACE and begins the logon procedure. This process consists of initializing all OWA related cookies and redirecting the client to the logon.aspx page. This redirection occurs via a HTTP LOCATION header. The ACE forwards the message, replacing the CAS role IP address with the CAS VIP address.



Note The CAS role is aware of the SSL-offload functionality of the ACE. To configure support for SSL-offloading on a CAS role, refer to:
<http://technet.microsoft.com/en-us/library/bb885060.aspx>

- Step 4** The client requests the logon page.
- Step 5** To verify the client supports Javascript, a heavily leveraged feature in OWA, the CAS role requests the client perform a Javascript request to receive the logon page with a *replaceCurrent* value equal to 1. If the client successfully sends this Javascript redirect the OWA server believes the client meets the necessary browser functionality.
- Step 6** The client requests a page through Javascript.
- Step 7** The OWA server returns the logon form.
- Step 8** The client posts the logon form to the **owaauth.dll**.
- Step 9** The CAS role authenticates the user via Active Directory and accepts the cookie set by ACE. The cookie information is then returned to the client in the Header of the HTTP response message. The following output shows the output of a single sessionid cookie being learned by the ACE:

```
show sticky database static
sticky group : sticky-cookie-group
type        : HTTP-COOKIE
timeout     : 20          timeout-activeconns : FALSE
  sticky-entry      rserver-instance      time-to-expire flags
-----+-----+-----+-----+-----+
2892825558555796692  CAS1:80                      never      -
```



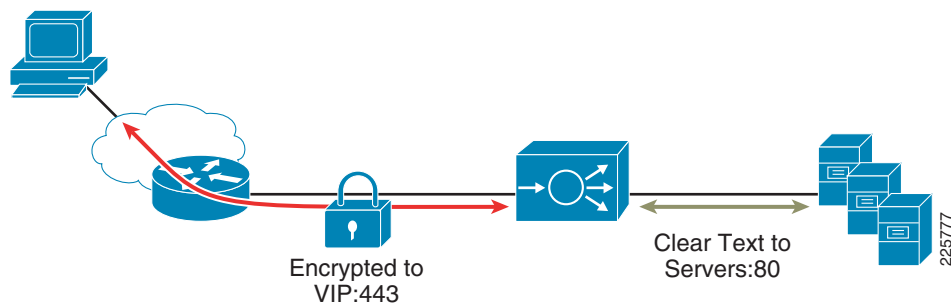
Note Microsoft recommends cookie or source IP-based session persistence. The ACE supports both in addition to cookie learning. Cookie learning allows the ACE to learn a cookie being applied by the CAS server itself.

- Step 10** The client redirects to the new location, ironically the same URL as the original request.
- Step 11** The OWA server responds with the client's home page.

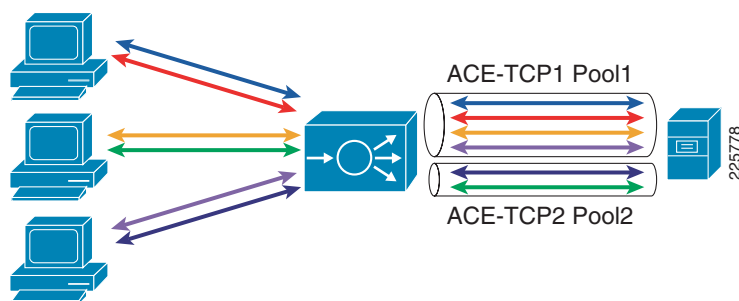
ACE Optimizations for OWA

The SSL Offload in conjunction with TCP Reuse capabilities of the Cisco ACE Module allow Enterprise MS Exchange environments to optimize their existing infrastructure by improving application performance by offloading compute intensive operations onto the network from the CAS servers themselves.

SSL termination occurs when the ACE, acting as an SSL proxy server, terminates a secure TCP connection from a client and then establishes a TCP connection to a CAS server. When the ACE terminates the SSL connection, it decrypts the ciphertext from the client and transmits the data as clear text back to the CAS Server listening on port 80. [Figure 27](#) depicts this process.

Figure 27 TCP Connection**TCP Reuse**

TCP server reuse allows the ACE to reduce the number of open connections on a CAS server by allowing connections to persist and be reused by multiple client connections. The ACE maintains a pool of TCP connections based on TCP options. New client connections can reuse those connections in the pool provided that the new client connections and prior server connections share the same TCP options. Figure 28 depicts how the feature works.

Figure 28 TCP Options

To configure TCP server reuse, use the **server-conn reuse** command in HTTP parameter-map configuration mode. The syntax of this command is as follows:

```
parameter-map type http TCP-REUSE
#Enable TCP reuse under the advanced parameter-map
server-conn reuse

policy-map multi-match OWA-POLICY-MAP
class OWA-VIP-443
#Apply policy map to VIP listening on port 443
appl-parameter http advanced-options TCP-REUSE
```

Conclusions

Microsoft CAS Servers can take advantage of Cisco's ACE Layer 7 load-balancing features SSL Offload and TCP Reuse thereby decreasing the CPU utilization, reducing the number of TCP connections, and increasing the number of HTTP requests being served by the CAS.

Lab validation for SSL Offload and TCP Reuse was done by simulating 100 users logging in and out for a period of 10 minutes and capturing measurements. HP Loadrunner 9.1 was connected to the Internet Simulation Site to perform the simulation. The first verification was done with the ACE performing Layer-4 loadbalancing and SSL termination taking place on the CAS Servers themselves. The second

verification applied the Layer 7 loadbalancing definition on the ACE to include SSL offload (see above). The third verification point configured TCP Reuse in an additive fashion in order to determine the incremental benefits of this feature.

For more information on SSL Offload and TCP Offload benefits, see the following whitepaper:

www.cisco.com/web/SG/learning/mindef/files/ACE_TCP_Offload.pdf

ACE for Outlook Anywhere

Outlook Anywhere allows the thick client application Outlook to connect to the CAS roles through HTTP/HTTPS. This communication technique was formerly known as RPC over HTTP(s). Outlook Anywhere essentially tunnels RPC traffic over HTTP or HTTPS. Outlook Anywhere allows firewalls without RPC fix-up capabilities to provide access to the Outlook client using the well-known ports of 80 and 443. In addition, Outlook Anywhere allows for MAPI connectivity over networks where latency is an issue. Traditionally, MAPI communication has issues where the network latency exceeds 250 milliseconds. Tunneling this traffic over HTTP(s) resolves this problem.

The ACE is more than capable of supporting HTTP and HTTPS traffic and is well positioned to provide services to Outlook Anywhere enabled clients. Outlook is not a web browser. Outlook does not support cookies. In this light, Microsoft recommends the use of source IP-based load balancing when leveraging a load-balancer with Outlook Anywhere clients. In addition, it is highly recommended to use secure transport of RPC, (i.e., SSL).

The configuration is identical to the OWA ACE configuration except for the type of session persistence. For Outlook Anywhere, deploy a source IP-based sticky group. To configure source IP-based sticky using the full IP address of the client, use the following ACE commands:

```
sticky ip-netmask 255.255.255.255 address source SRC-STCKY-GRP
serverfarm CAS_FARM
```

Security Considerations for the CAS Role

The CAS role is an access point for users to reach their mailbox. As such, the protection of the CAS role itself and its data paths is a primary consideration when planning an Exchange Server 2007 environment. Microsoft recommends the use of secure transport for all CAS communications: it is the default configuration of this server role. The ACE allows the network administrator to configure a central PKI infrastructure on the module and offload this client access service from the CAS roles. The ACE simplifies certificate management, preserves server resources and provides the level of transport security the Exchange environment demands.

From a traditional firewall perspective, the CAS role must be able to access the ports listed in [Table 3](#), to provide connectivity to all external clients and other Exchange entities. Configure the firewall device to provide the appropriate ACLs to support this configuration. For more information on the deployment of security policies for the Cisco ASA, refer to the following URL:

<http://www.cisco.com/en/US/docs/security/asa/asa80/configuration/guide/inspect.html>

Note that some of these features may be disabled by the server administrators and therefore will not require a "hole" in the firewall.

Table 3 CAS Communication Ports *msxchng_cmcvmc.mif*

Communication	Port(s)
Autodiscover service	80/TCP, 443/TCP (SSL)
Availability service	80/TCP, 443/TCP (SSL)
Outlook Web Access	80/TCP, 443/TCP (SSL)
POP3	110/TCP (TLS), 995/TCP (SSL)
IMAP4	143/TCP (TLS), 993/TCP (SSL)
Outlook Anywhere (formerly known as RPC over HTTP)	80/TCP, 443/TCP (SSL)
Exchange ActiveSync application	80/TCP, 443/TCP (SSL)
Client Access server to Unified Messaging server	5060/TCP, 5061/TCP, 5062/TCP, a dynamic port
Client Access server to a Mailbox server that is running an earlier version of Exchange Server	80/TCP, 443/TCP (SSL)
Client Access server to Exchange 2007 Mailbox server	RPC. (Dynamic Ports)
Client Access server to Client Access server (Exchange ActiveSync)	80/TCP, 443/TCP (SSL)
Client Access server to Client Access server (Outlook Web Access)	80/TCP, 443/TCP (SSL)
WebDAV	80/TCP, 443/TCP (SSL)

**Note**

Microsoft strongly recommends the use of an application layer firewall to provide protection up the TCP stack.

**Note**

For more information, refer to the *Data Path Security Reference* document at the following URL <http://technet.microsoft.com/en-us/library/bb331973.aspx>. This document is a comprehensive guide to all ports and data paths leveraged by the Microsoft Exchange 2007 server roles and clients.

Route Health Injection

Route Health Injection (RHI) allows the ACE to advertise host routes associated with any number of active virtual IP addresses hosted by the device. The injection of the host route to the remaining network offers Layer 3 availability and convergence capabilities to the application environment. In the Exchange test bed, the ACE advertises the VIP frontending the CAS serverfarm to the Multilayer Switch Feature Card (MSFC) routing table. The MSFC distributes the route into the IGP as a host route.

The following configuration is necessary on the ACE module to enable RHI:

```
#VIP 10.5.103.11 listening on port 443
class-map match-all OWA-VIP-443
  2 match virtual-address 10.5.103.11 tcp eq https

policy-map multi-match OWA-POLICY-MAP
  class OWA-VIP-443
    loadbalance vip inservice
    loadbalance policy OWA-LB-POLICY
#RHI will advertise 10.5.103.11 as host route when active
  loadbalance vip advertise active
```

The host route is present on the MSFC:

```
DCAAL-DC1-Agg-1#sh ip route 10.5.103.11
Routing entry for 10.5.103.11/32
  Known via "static", distance 77, metric 0
  Routing Descriptor Blocks:
    * 10.5.103.4, via Vlan103
      Route metric is 0, traffic share count is 1
```

Route Health Injection deployments in the data center are described in Chapter 5 of the *Data Center—Site Selection for Business Continuity*:

http://www.cisco.com/en/US/netsol/ns656/networking_solutions_design_guidances_list.html#anchor3

Cisco WAAS Deployment for the CAS Role

WAAS Traffic Interception

Cisco WAAS supports both out-of-path and in-path deployment mechanisms for the Wide Area Engine (WAE). The following out-of-path deployment mechanisms are all able to place the WAE logically in-path but physically off-path:

- WCCPv2
- PBR
- ACE

WCCPv2 is the preferred off-path interception mechanism for WAAS. WCCP with WAAS is currently supported on a variety of routing platforms, including the Integrated Services Router (ISR models 1800, 2800, and 3800), 3700 series Access Routers, Cisco 7200 series routers (with NPE-400, NPE-G1, NPE-G2 only), 7600 routers, and ASR 1000 series routers. WCCP is also supported on a variety of switching products, including the Catalyst 3560/3750, Catalyst 4500/4948, and Catalyst 6500.

Inline interception

Cisco WAAS supports a physical in-path option, which allows the WAE to be inserted physically between two network devices such as the branch office LAN switch and branch office WAN router. The Cisco WAAS inline card has 4 10/100/1000BaseT Ethernet ports in two port groups. Each port group provides a fail-to-wire bypass service with mechanical relays to ensure that network connectivity is not interrupted should a device fail or a software crash be encountered on the WAE. The Cisco WAAS in-path card is supported in any Cisco WAAS appliance model.

Both methods of traffic interception are utilized in this solution. Inline interception is used for WAE's that optimize storage replication traffic and a hybrid model is used for branch to data center traffic. In this scenario the branch WAE is configured inline for simplicity and performance and the data center the WAE supporting branch application traffic is configured for WCCP interception to provide flexibility and ability to select which traffic to be optimized.

WAN Optimization Traffic Patterns

As previously defined, the CAS role provides access to the Microsoft Exchange environment via all non-MAPI forms of communication. These different types of client applications and protocols involved are detailed in [Table 4](#).

Table 4 **Client Access Server Role Communication**

Client Type	Communication Protocol
IMAP Client	IMAP4 / SMTP
POP Client	POP/ SMTP
Outlook Voice Access	RTP
Exchange ActiveSync	HTTP/HTTPS
Outlook Anywhere	MAPI over RPC over HTTP/HTTPS
Outlook Web Access (OWA)	HTTP/HTTPS
Outlook	MAPI over RPC

OWA allows users to have a robust or "thick-client" application experience within a web browser. Microsoft continues to introduce new features and functionality to this product that has standardized on the HTTP protocol for conducting remote transactions. Cisco technologies such as ACE and WAAS are well positioned to provide network-based services in an OWA environment because Microsoft uses this standard form of communication.

Outlook Anywhere allows the Outlook client application to access its designated mailbox server using RPC over HTTP/HTTPS. This technology provides efficient access to users through enterprise firewall deployments. The CAS role supports this functionality, as do the Cisco ACE and WAAS technologies.

The following section details the flow of OWA or Outlook Anywhere traffic from an enterprise branch to the CAS roles residing in the enterprise edge of the data center. These traffic patterns include:

- [Egress Client to Enterprise Edge, page 62](#)
- [Ingress Enterprise Edge to CAS Role, page 63](#)
- [Egress CAS Role to Enterprise Edge, page 65](#)
- [Egress Enterprise Edge to Client, page 67](#)

This section details the transparent integration of WAAS technology.

Egress Client to Enterprise Edge

Clients in the branch using OWA or Outlook Anywhere services should consider using Cisco's application acceleration technologies, such as Cisco WAAS. Cisco WAAS requires a minimum of two WAE devices to auto-discover and deliver applicable application optimizations. To leverage these transparent optimizations across the WAN, deploy one or more WAEs at the remote branch and one or more WAEs at the enterprise data center, depending on availability and scalability requirements.



Note

For more information on Cisco WAE branch deployments, see the *Enterprise Branch Wide Area Application Services Design Guide* at the following URL:
<http://www.cisco.com/en/US/docs/solutions/Enterprise/Branch/WAASBr11.html>

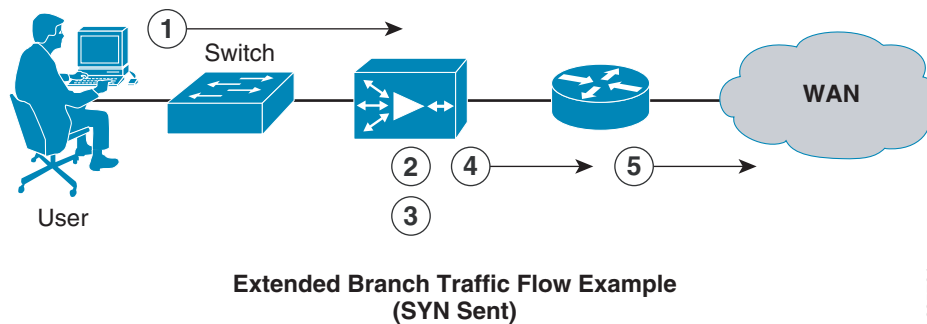
Figure 29 Example Branch Traffic Pattern

Figure 29 depicts the traffic pattern an OWA or Outlook 2007 client used to reach the Microsoft Exchange 2007 deployment in the data center, specifically the CAS role residing at the enterprise edge. The users must cross the WAN, a potential application bottleneck with its inherent latency and bandwidth challenges. To address this issue, the Cisco WAE is deployed at the remote branch. The following steps describe the initiation of communication between the Exchange client and the CAS role in the data center:

- Step 1** The OWA or Outlook Anywhere client initiates a TCP connection via SYN packet to the ACE VIP frontending the CAS server farm. The VIP is advertised through RHI from the ACE.
- Step 2** The WAE transparently intercepts the TCP SYN using its inline interface and sends it up to the device CPU for the optimization process to be applied.
- Step 3** The WAE optimization process applies a new TCP option (0x21) to the packet if the application is identified for optimization by an application classifier. The WAE adds its device ID and application policy support to the new TCP option field. This option is examined and understood by other WAEs in the path as the ID and policy fields of the initial WAE device. The initial ID and policy fields are not altered by another WAE.

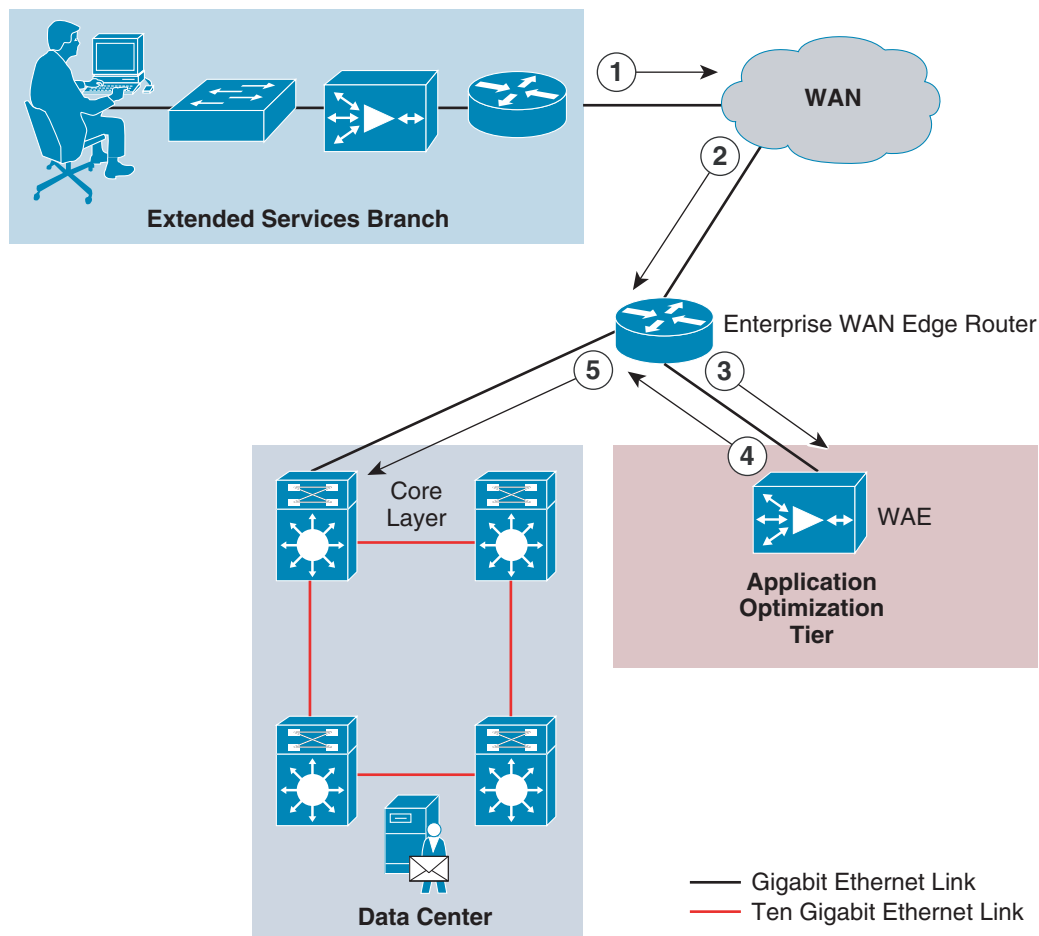


Note The Cisco WAE device has default optimizations enabled for HTTP/HTTPS or web traffic. Encrypted traffic will not benefit from all of the optimization techniques offered by the ACE. Compression and data redundancy elimination optimizations will be greatly reduced if not eliminated when encrypted traffic is processed by the WAE. However, the transport flow optimizations of the Cisco WAAS solution will continue to offer benefits including large initial TCP window sizes, scaling, and congestion handling techniques to reduce data loss.

- Step 4** The WAE forwards the packet to the branch router that is configured as the default gateway for the WAE devices. Note that the source and destination IP addresses of the initial request of the remote client remain unchanged.
- Step 5** The router forwards the packet to the CAS VIP address across the WAN.

Ingress Enterprise Edge to CAS Role

Ingress traffic across the WAN destined for the Exchange CAS server farm will cross the WAN edge routers. Figure 30 clarifies the subsequent actions that occur for application optimization of the OWA or Outlook Anywhere clients.

Figure 30 *Ingress Traffic from the WAN to the Data Center*

In Figure 30, the following occurs:

- Step 1** Traffic exits the branch with the WAE TCP option (0x21) set.
- Step 2** Traffic enters the WAN edge router with a destination IP address of the CAS VIP. The router forwards the packet to an application optimization-tier hosting a WAE farm via WCCPv2. The following commands are necessary on the WAN Edge Router to enable WCCPv2 redirection:

```
#Enable the WCCP service.
ip wccp 61
interface FastEthernet2/8
description <<*** Interface to WAN ***>
ip address 10.5.200.1 255.255.255.0
ip wccp 61 redirect in
!
interface FastEthernet2/1
description <<*** Interface to WAE ***>
ip address 20.20.20.1 255.255.255.0
```

The **ip wccp 61** service group command instructs the router to inspect traffic for any TCP packets. Matching packets should be load-balanced amongst service-group attached WAEs using assignment based on the source IP address of the packet. Inbound redirection combined with CEF is recommended to reduce the resource utilization of the router.

- Step 3** The dedicated WAE VLAN carries the redirected traffic from the router to a specific WAE via MAC rewrite. The WAE receives the SYN containing the TCP option introduced by the branch WAE. The local data center WAE is aware of the branch WAE and what optimizations it supports. The WAE appends its ID and application policy support to the existing TCP option 0x21 field while storing the ID and application policy of the remote branch WAE. The WAE is registered to the WCCP service group in the router via the following commands:

```
#Register the WAE with the WCCP service groups in the WAN Edge Router
wccp router-list 1 <IP Address of WAN Edge Router>

#The routers in the list will use L2 redirection (MAC address rewrite) to send interesting
#traffic to the WAE
wccp tcp-promiscuous router-list-num 1 l2-redirect assign-method-strict
wccp version 2
```

- Step 4** The WAE returns the packet to the router based on the egress method configured on the WAE, in this case via GRE. The WCCP protocol allows for the auto-negotiation of the GRE tunnel. The following command enables the GRE functionality:

```
egress-method negotiated-return intercept-method wccp
```



Note

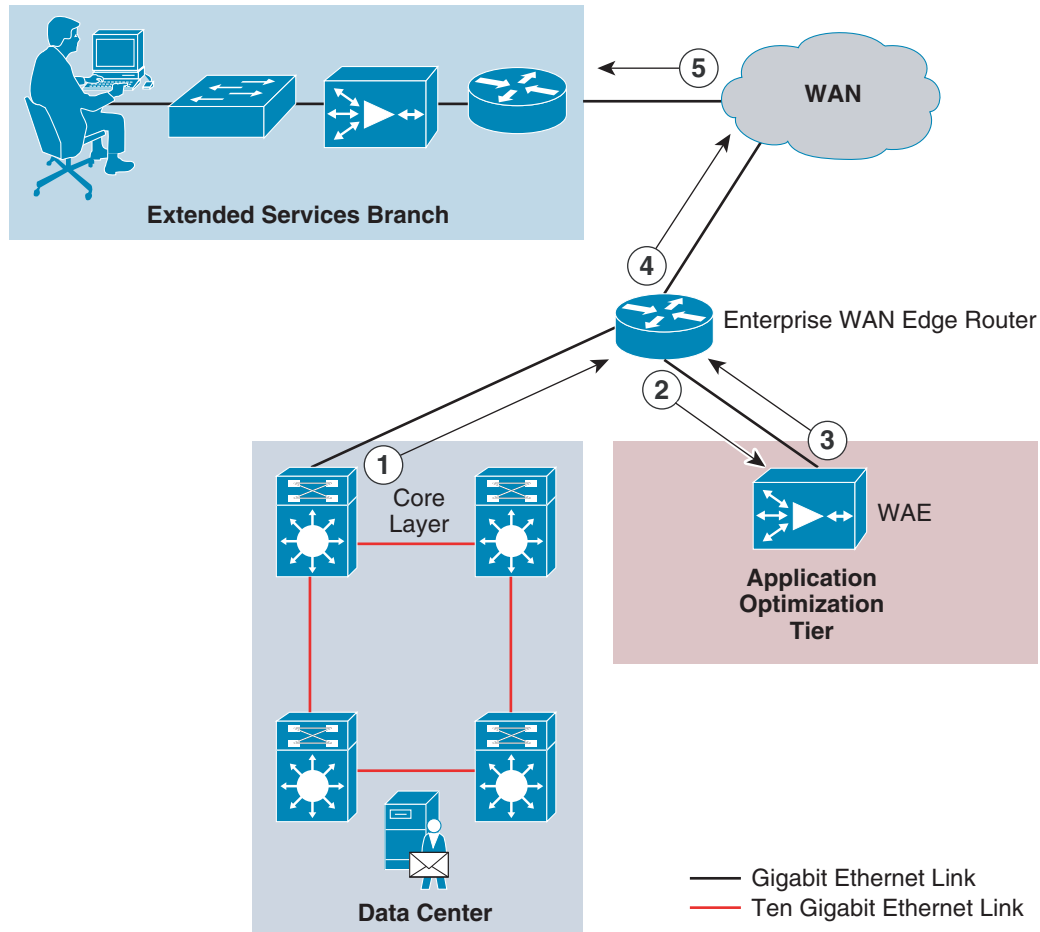
The default form of WAE egress traffic forwarding is IP-based. Typically, one uses IP forwarding when the WAE farms have a dedicated subnet, where as GRE based forwarding is leveraged when the WAE farm resides within the client and or server subnet. Use the **show egress-method** command on a WAE device to review the policy.

- Step 5** The packet is routed to the ACE CAS VIP through Layer 3 for ACE-to-CAS traffic flows.

Egress CAS Role to Enterprise Edge

One of the primary tenets to a well-designed network infrastructure is the ability to control traffic flow across the enterprise. From branch-to-data center or data center-to-data center, predictable traffic patterns are a hallmark to the stability and availability of an application. Traffic returning from the enterprise edge CAS roles with a symmetric path out of the data center will benefit from the application optimization offered at the WAN edge, providing acceleration to the HTTP/HTTPS transport of OWA and Outlook Anywhere clients.

[Figure 31](#) defines the flow of traffic from the data center core to the branch in the test bed's Exchange 2007 environment. It is important to recognize that the traffic returning from the core is directed to the same WAE on egress as on ingress to enable WAN optimization.

Figure 31 **Return CAS Traffic to the WAN Edge**

As shown in [Figure 31](#), the following flow occurs from core to branch:

- Step 1** SYN/ACK returned from the CAS serverfarm is routed to the WAN edge router. The router receives the packet on its internal interface with WCCPv2 enabled, service group 62. The following commands are necessary to enable this service on the 7200 router platform.

```
#Enable WCCP service group
ip wccp 62
!
#Apply the service to the ingress interface from the data center core
interface GigabitEthernet0/0
description DC1 Core 1
ip address 10.5.11.2 255.255.255.0
ip wccp 61 redirect in
!
interface GigabitEthernet0/1
description DC1 Core 2
ip address 10.5.10.2 255.255.255.0
ip wccp 62 redirect in
```

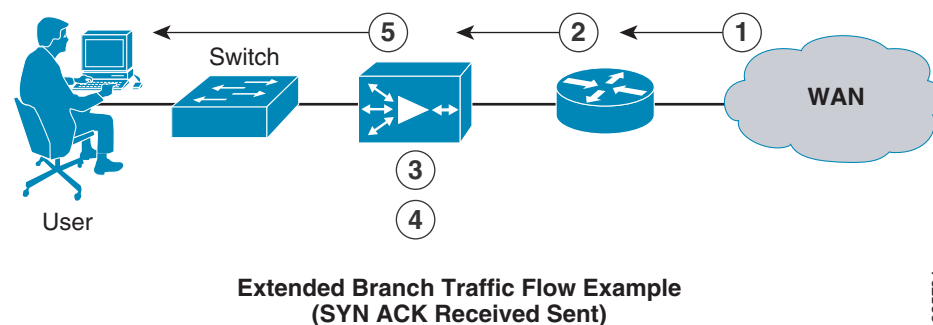
The **ip wccp 62** service group command instructs the router to inspect traffic for any TCP packets. Matching packets should be load-balanced amongst service-group attached WAEs using assignment based on the destination IP address of the packet.

- Step 2** The packet returns to the same WAE, maintaining symmetric traffic flow and WAN optimization consistency.
- Step 3** The WAE receives the SYN/ACK and adds TCP option 0x21 to the packet with its ID and application policy support. The packet is returned via the egress-method policy negotiated via WCCP to the intercepting router.
- HTTP/HTTPS traffic is optimized by default on the Cisco WAE platform. OWA and Outlook Anywhere traffic fall in this category.
- Step 4** The packet is routed to the branch.
- Step 5** The branch router receives the packet.

Egress Enterprise Edge to Client

Figure 32 depicts the return traffic for TCP flows coming from the data center. The return of traffic to the WAE device located in the branch will complete the auto-negotiation and begin WAN optimization for the OWA and Outlook Anywhere clients.

Figure 32 Returning Traffic from the Data Center to Branch Client



The flow in Figure 32 shows the following:

- Step 1** The branch router receives the packet on the interface with a WCCP service group defined for inbound traffic interception.
- ```
interface GigabitEthernet0/0
 description ** WAN interface **
 ip wccp 62 redirect in
```
- Step 2** Traffic is forwarded to the interface supporting the client subnet.
- Step 3** The WAE device intercepts the TCP flow and sends it up to the CPU optimization process.
- Step 4** The WAE is aware of the WAE in the data center because the SYN/ACK TCP option 0x21 contains an ID and application policy. The auto-negotiation of the policy occurs as the branch WAE compares its application-specific policy to that of its remote peer defined in the TCP option. At this point, the data center WAE and branch WAE have determined the application optimizations to apply on this specific TCP flow.

The **show statistics tfo connection** command on the branch or data center WAEs details the results of auto-negotiation for each TCP flow. Below is an example output for an SSL connection between an OWA SSL client and the CAS VIP. This view shows that only TCP optimizations are occurring for the HTTPS transaction, the four Ts indicate that TFO optimization is occurring at the local and remote WAE.

```

Optimized Connection List
Policy summary order: Our's, Peer's, Negotiated, Applied
F: Full optimization, D: DRE only, L: LZ Compression, T: TCP Optimization
Local-IP:Port Remote-IP:Port ConId PeerId Policy
10.5.201.50:28073 10.5.103.11:443 17174 00:14:5e:a4:4f:66 T,T,T,T

```

**Step 5** The WAE forwards the traffic to the switch supporting the client.

---

## Cisco WAAS Deployment for EMC RecoverPoint

This section discusses the integration of the Cisco WAAS Replication Acceleration mode with EMC RecoverPoint. Data replication appliances such as EMC's RecoverPoint are dependant on the IP network infrastructure for their performance. Wide Area Networks (WANs) introduce latency, packet loss, congestion and bandwidth limitations that impact data replication performance. Additionally, WAN links are inefficiently used, having to transmit repeated data patterns due to the similarity of data in an organization. With limited replication/backup time windows, IT organization struggle to process the ever increasing amounts of data that need to be replicated across their WAN links.

Data replication traffic has unique characteristics that require dedicated and specialized resources to expedite processing and overcome the challenges and limitations imposed by the WAN. Data replication traffic is by nature high in volume with a low number of TCP connections. These connections are long lived per session and they are persistent over a long time. This is unlike branch office to data center traffic which is low volume with a high number of TCP connections that are short lived per session. A solution to overcome these limitations and allow efficient backup within the allocated time is needed.

Cisco WAAS Version 4.0.19 introduces a new optimization mode called **Replication Accelerator** mode which is designed to accelerate the storage systems' replication activities between data centers (also referred to as DC-to-DC). When a Cisco WAE (WAE-7371 or WAE-7341) is enabled for **Replication Accelerator** mode the default policy will be automatically changed such that the WAE will accelerate only traffic related to data replication operations. As soon as **Replication Accelerator** mode is enabled the WAEs will be able to detect each other without any additional configuration. In addition to auto discovery and appropriate policy changes the TCP buffers will also be tuned automatically to support typical data replication traffic. Thus, the installation and configuration of WAEs for replication acceleration have become extremely easy with Cisco WAAS. In this mode, the disk I/O performance has been fine tuned to support the need of typical data center applications. Similarly, the TCP and DRE optimization parameters are also adjusted to suit the data center applications.

### Deployment Scenarios

As in a branch office-to-data center (application acceleration) scenario, traffic can be intercepted and/or redirected to a WAE that is enabled for replication acceleration. This can be done using one of the following methods:

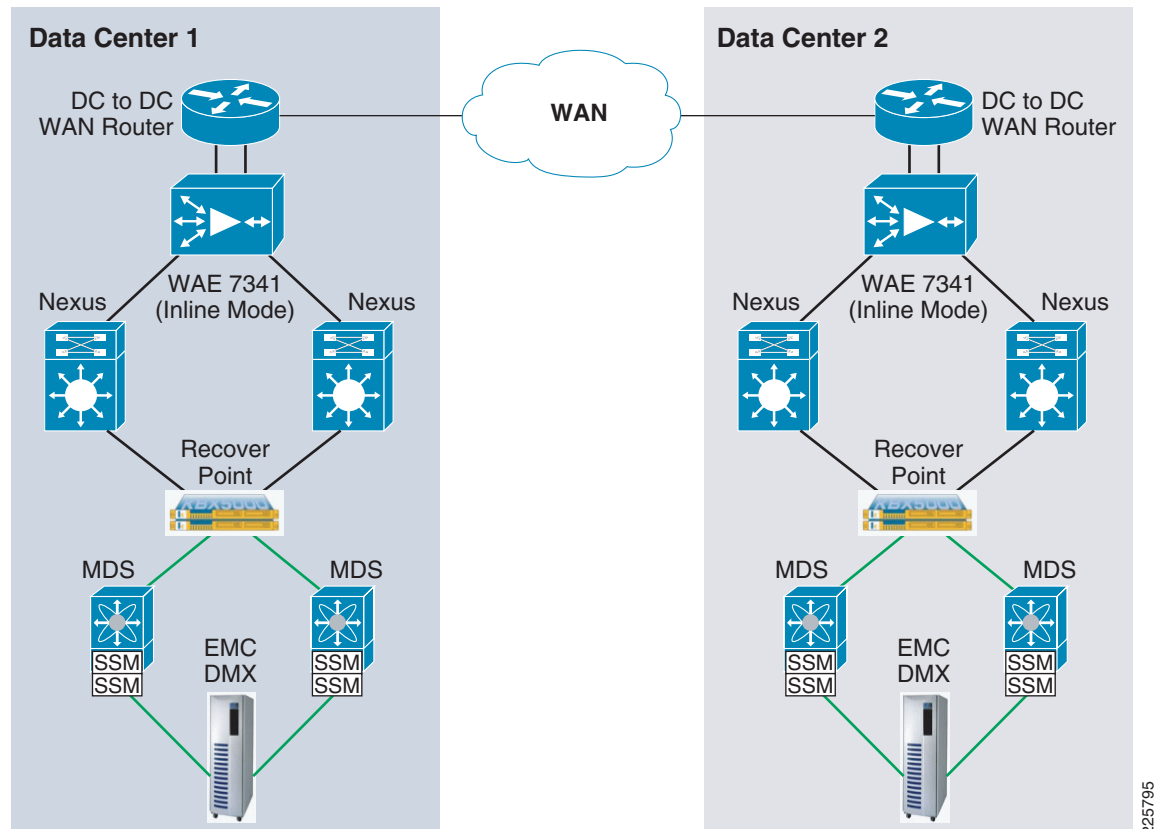
- Inline (preferred for replication acceleration)
- WCCP (N+1 cluster for high availability and load balancing)
- ACE, CSM module, CSS, or similar device

This solution uses the recommended approach on inline mode. The following section describes the procedure for configuring the WAE for replication acceleration—inline mode. In the scenario below WAE was installed in **Inline** and enabled for **Replication Accelerator** mode. To obtain an optimal performance for data replication, Cisco WAAS recommends using inline mode and installing the WAE that will be performing data replication optimization as close as possible to the storage systems. Thus, the WAE that is configured for **Replication Accelerator** will not have to receive non-Data Replication (DR) traffic. The inline mode also supports port-channel, which enables load-balancing and high-availability if one of the physical link fails.

**Note**

The procedure for configuring port channel on WAE can be found at the following URL  
[http://www.cisco.com/en/US/docs/app\\_ntwk\\_services/waas/waas/v401\\_v403/command/reference/glob\\_cfg.html](http://www.cisco.com/en/US/docs/app_ntwk_services/waas/waas/v401_v403/command/reference/glob_cfg.html) - wp5409468

**Figure 33** Cisco WAE in inline Mode optimizing Data Replication over Native IP



The following configuration example shows how to enable **Replication Accelerator** mode on WAE. In the scenario shown in [Figure 33](#) **Replication Accelerator** mode must be enabled on WAEs at both sides.

```
#Enable Replication Acceleration on WAE
device mode replication-accelerator
!
#Management interface for Central Manager
interface GigabitEthernet 1/0
 ip address 172.28.210.80 255.255.255.0
 exit
!
interface InlineGroup 1/0
 inline vlan all
 exit
!
interface InlineGroup 1/1
 inline vlan all
 exit
```



**Note**

After the WAE is configured for replication acceleration, it must be reloaded for the changes to take effect. As soon as it is reloaded, the default policy will be automatically changed such that the WAE engine will accelerate only traffic related to data replication operations.

**Management**

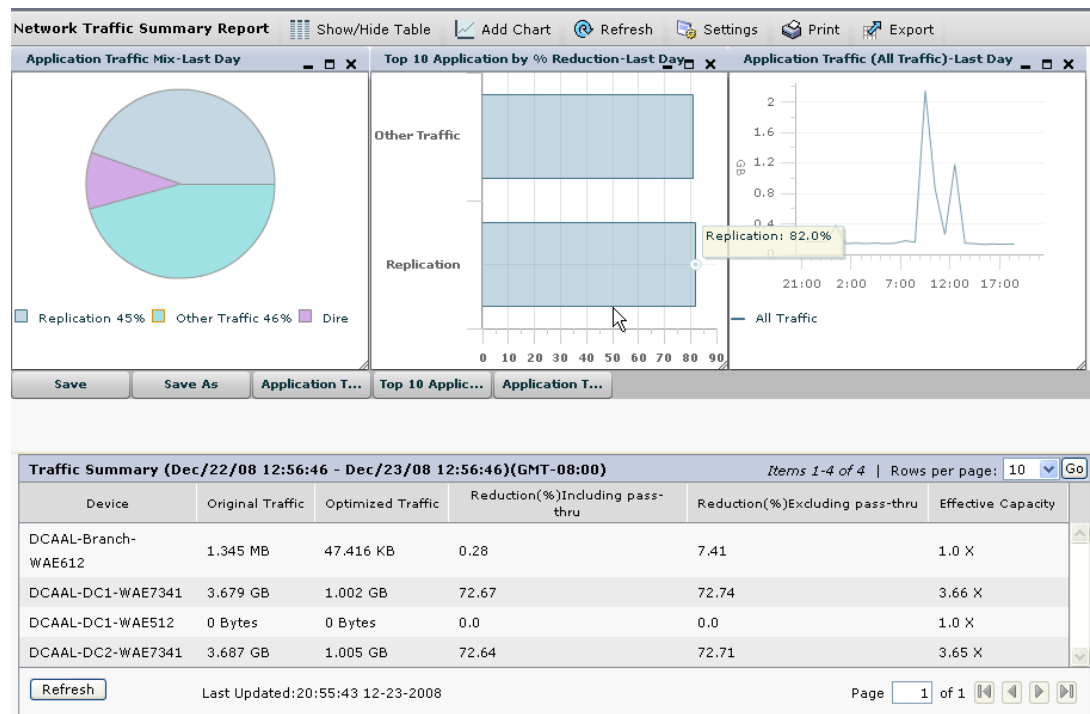
WAAS Central Manager is used to manage both **Replication Acceleration** mode appliances as well as Application Acceleration appliances. Although RA mode devices use WAAS Release 4.0.19 code base and AA mode devices typically will be running the newer Release 4.1.1 code, the Central Manager can support both platforms simultaneously provided that it is running Release 4.1.1.

**Limitations/Restrictions**

- **Replication-Accelerator** mode is supported only on WAE-7371 and WAE-7341 running Release 4.0.19.
- Device mode **Replication-Accelerator** can only be configured through CLI.
- As the auto-discovery mechanism distinguishes between WAE running **Replication Accelerator** mode versus the WAE running **Application Accelerator** mode the connections between two WAEs running incompatible device mode will not be optimized. Instead, these connections will be established as pass-through connections.
- Maximum supported DRE fanout is 9 peers using WAE-7371 and 4 DRE peers using WAE-7341.
- On WAE-7371, maximum supported concurrent TCP connections are 5000. Similarly, on WAE-7341 maximum supported concurrent TCP connections are 2500.
- When a WAEs device mode change from **Application Accelerator** to **Replication Accelerator** the default policy will be changed as well. So, any changes to existing policy while WAE was in **Application Accelerator** mode will be lost after the new device mode becomes effect. The same is applicable if a WAEs device mode changes from Replication Accelerator to **Application Accelerator**.

**WAAS Performance Benefits for RecoverPoint**

During the solution validation process for this project, basic performance tests were performed to determine relative performance gains for RecoverPoint when used in conjunction with Cisco WAAS. Two test scenarios were performed, one using a large file transfer of approximately 10Gb which was copied from local storage to a replicated storage group attached to the same server. The second test case used Exchange Load Gen tool to create volume email traffic directed to the primary mail box server in data center 1 which was then replicated to the secondary mail box server in data center 2. See [Figure 34](#).

**Figure 34** WAAS Performance for RecoverPoint

WAAS Optimization for RecoverPoint results in approximately 73 percent reduction in traffic across the WAN or a 3.66 times increase in effective capacity. These results are provided as a reference for what can be expected. Results will vary depending on the data being replicated across the WAN. In the testing scenarios, some transactions are repeated which will increase WAAS ability to employ DRE caching. In the real world, scenarios some data may be more unique with respect to email content, but when replicating Exchange log files there will be much repeated data as RecoverPoint replicated the entire file each time it changes and by nature only the tailend of the file has changed.

## Cisco IronPort Secure Email Appliance Deployment

This section discusses the integration of the Cisco IronPort Secure Email Appliance with Microsoft Exchange 2007. As mentioned earlier, the Cisco IronPort C Series is being deployed as a secure email appliance and also as an SMTP SmartHost relay. The following subsection discusses the configurations to setup outbound and inbound SMTP messaging between the HT and IronPort.

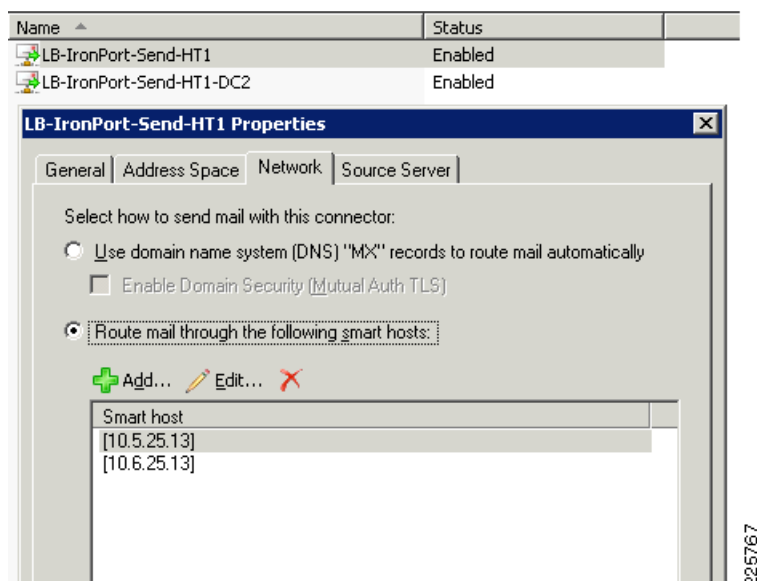


### Note

The messaging security features of the IronPort are too many to include here and their unique settings are customer implementation specific and therefore, are not covered. Refer to the Cisco IronPort reference links in this document to find out more on deploying the messaging security features of the Cisco IronPort C Series appliance.

### Outbound SMTP

The SMTP SmartHost configuration between the HT and the IronPort is straightforward. Figure 35 shows the Exchange HT screen output for the Send Connectors (outbound). In this example, there are two send connectors; one from each HT (HT in DC1 and HT in DC2). Both Send Connectors send mail to the IronPort appliances through the smart host configuration. SMTP mail is processed equally across the connectors between the IronPort and HT role.

**Figure 35** Exchange 2007 Hub Transport Send Connectors

The HT also needs to have the Remote Domains configured at the Exchange Organization level. Most often, the Remote Domain is defined as "\*", indicated 'All' domains.

Figure 36 shows the IronPort side of the configuration. The two HTs (10.5.101.50 and 10.6.201.50) are authorized senders within the IronPort configuration.

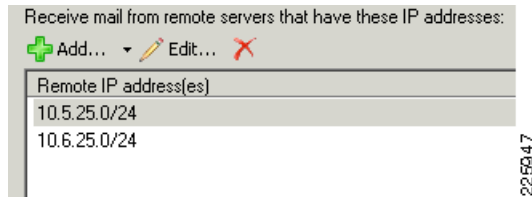
**Figure 36** IronPort Configuration

| Sender      | Comment |
|-------------|---------|
| 10.6.201.50 | HT1-DC2 |
| 10.5.101.50 | HT1-DC1 |

These two configurations complete the outbound mail portion of the pair – HT has outbound SMTP SmartHost defined (Figure 35) and the IronPort is allowing Senders (Figure 36) in its Host Access Table (HAT).

### Inbound SMTP

The HTs need to have the Receive Connectors (inbound) defined as well as the Accepted Domains. The HT has one Receive Connector defined for the two address ranges that the IronPort appliances reside on in both data centers. Figure 37 shows that the HT will receive mail from servers (IronPort appliances) in the 10.5.25.0/24 (Data Center 1) and 10.6.25.0/24 (Data Center 2). Alternatively, the specific IP address of the IronPort could be defined to nail this down further.

**Figure 37** Hub Transport Receive Connector - Accepted IP Addresses

The Accepted Domain is shown in Figure 38.

**Figure 38** Hub Transport Accepted Domain

| Name    | Accepted Domain | Type          | Default |
|---------|-----------------|---------------|---------|
| ese.com | ese.com         | Authoritative | True    |

The Ironport needs to have the approved domains listed in its SMTP Routes configuration along with the destination hosts to relay the SMTP mail to (the HTs). Figure 39 shows the SMTP routes for ese.com, all sub-domains within ese.com (.ese.com) and destination hosts defined for both HTs:

**Figure 39** IronPort Approved Domains

| Receiving Domain | Destination Hosts        |
|------------------|--------------------------|
| .ese.com         | 10.5.101.50, 10.6.201.50 |
| ese.com          | 10.5.101.50, 10.6.201.50 |

**Note**

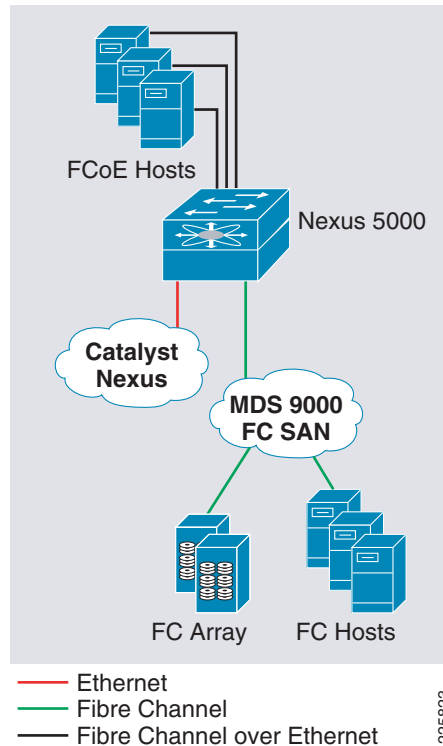
In addition to TCP port 25, there will be other ports that need to be opened on the firewall. Refer to the IronPort C Series “Basic\_User\_Guide” for the relevant TCP/UDP ports that may need to be opened.

**FCoE**

The solution architecture uses Cisco® Nexus 5000 Series Switches, part of the Cisco Nexus family of data center-class switches, to deliver a unified fabric at the network access layer or edge, where servers connect to the LAN and SAN. To integrate with existing infrastructure, the Cisco Nexus 5000 Series provides native FibreChannel uplinks to facilitate connection with installed SANs and available SAN switches. Other Cisco products, including the Cisco MDS 9000 family, will support FCoE in the future to extend consolidation beyond the access layer.

The unified fabric tested in this solution architecture consists of Nexus 5000s with N5K-M1404 modules. The FibreChannel interfaces connected with two 4Gb FibreChannel ports configured in a port channel back to MDS in the fabric. A 10-Gigabit SFP+ is connected to the access layer switch for Ethernet connectivity. ESX hosts with Qlogic QLE8042 CNA adapters are connected to the Nexus 5000 with Copper Twinax cables.

Unified fabric is then configured between the ESX hosts and the Nexus 5000 as shown in as shown in Figure 40.

**Figure 40**      **Nexus 5000 Unified Fabric Connectivity**

The main benefits of the topology in [Figure 40](#) include the following:

- Consolidation of server I/O and cables on a single interface whereas the traditional configuration included multiple Ethernet and FibreChannel interfaces and cables.
- Significantly improved energy efficiency due to fewer interfaces and adapters, resulting in less power required per server.
- More affordable model for more servers to gain access to the SAN for the first time.
- FibreChannel SAN investment protection: no changes to the SAN operating and management model; FCoE-enabled servers continue to access SAN-attached FibreChannel storage.

## VMware Design and Deployment

The following section discusses the tested layout and deployment of VMware in support of Exchange 2007. The areas of focus for the VMware deployment will be:

- VMware ESX
- VMware HA
- VMware DRS
- VMware SRM

## VMware ESX

VMware ESX Version 3.5 was used in the deployment of the fully virtualized Exchange 2007 environment. As [Figure 10](#) shows, there are two VMware vCenter servers; one vCenter server for each data center location. A total of four ESX hosts running at data center 1 (ESX 3-6) and another four ESX hosts (ESX 7-10) were running at data center 2. All Microsoft Active Directory, Windows Update, and Exchange 2007 servers were running on ESX VMs with Windows Server 2008 Data Center Edition installed. The entire ESX environment consisted of two ESX clusters. One cluster “DC1-CLS” had the four ESX hosts as members for data center 1 and another cluster “DC2-CLS” has the other four ESX hosts as members for data center 2. These clusters had VMware HA and DRS enabled (see [VMware HA, page 76](#) and [VMware DRS, page 77](#)).

### ESX Networking

The VMware environment between both data center sites were duplicated exactly. The following description of the network applies to data center 1 but is duplicated in data center 2 and only naming, VLAN numbers and IP addressing is what differs. This information is a high level summary as there is very little that is unique to Exchange 2007 and therefore, there is no reason to define in detail the entire configuration for either the Cisco components or the VMware configuration.

Each ESX host had multiple network connections to the data center access layer. Two of the ESX hosts had multiport 1Gbps NIC cards that were connected to the “dc1-acc-1” switch. Multiple VLANs were trunked to the ESX host to support the various VLANs used by each VM. The other two ESX hosts used the Converged Network Adapter (CNA) to provide unified 10Gbps network and SAN access. These hosts were connected to the Nexus 5020 “dc1-n5k-1”. The following VLANs were trunked into the ESX hosts via the 1Gbps attached “dc1-acc-1” on ESX hosts 4 and 5:

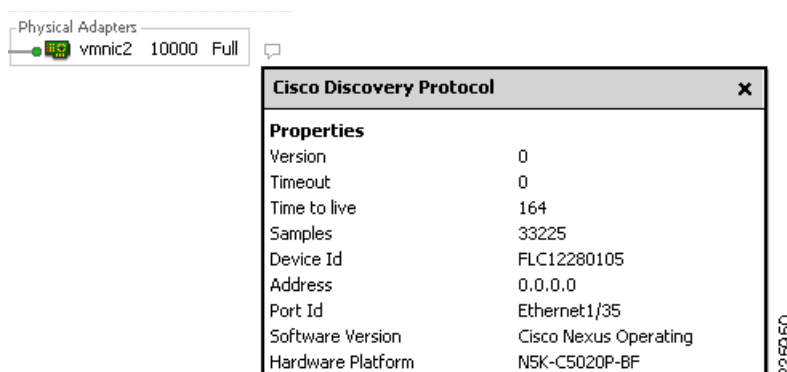
- VLAN100-AD-DNS
- VLAN101-HT
- VLAN102-MBX
- VLAN103-CAS

The following VLANs were trunked into the ESX hosts through the 10Gbps CNA attached “dc1-n5k-1” on ESX hosts 3 and 6:

- CNA-VLAN100
- CNA-VLAN101
- CNA-VLAN102
- CNA-VLAN103

The Nexus 5020 provides network access for the ESX hosts through multiple 10Gbps trunk links to the data center aggregation layer.

The VMware Infrastructure Client output in [Figure 41](#) shows the CNA adapter (vmnic2) connected to the Cisco Nexus 5020 switch with CDP configured between them.

**Figure 41 ESX Host Network Using CNA**

All ESX hosts had the same VLANs trunked to them so that VMware HA, DRS, and VMotion can function properly. The network deployment specifics for the Cisco infrastructure as well as the VMware ESX environment are based on Cisco and VMware best practices. More information regarding the best practices can be found in the [VMware Infrastructure in a Cisco Network Environment](#) and the [VMware ESX 3.5 U2](#) documentation.

### ESX SAN

Similar to the network configuration above, the layout is not specific to Exchange 2007. Two ESX hosts (ESX 4-5) were connected to a pair of Cisco MDS SAN switches through 4Gigabit FibreChannel (see [Figure 10](#)). The other two ESX hosts (ESX 3 and 6) were connected to the Nexus 5020 Unified I/O switch via the CNA. Again, the CNA offers unified connectivity of both the Ethernet network connection and the SAN connection. The Nexus 5020 provides SAN access for the ESX hosts via multiple 4Gigabit FibreChannel connections to the Cisco MDS fabric. The ESX storage adapter configuration shows the CNA ports as 4Gigabit FCoE HB, as illustrated in [Figure 42](#).

**Figure 42 ESX Host SAN Using CNA**

**Storage Adapters**

| Device                                        | Type          | SAN Identifier          |
|-----------------------------------------------|---------------|-------------------------|
| <b>ISP8432-based 4Gb FCoE PCI Express HBA</b> |               |                         |
| vmhba4                                        | Fibre Channel | 21:00:00:1b:32:0a:d1:be |
| vmhba5                                        | Fibre Channel | 21:01:00:1b:32:2a:d1:be |

All LUNs located on the EMC storage were mapped so that all ESX hosts had full access to them so that VMware HA, DRS, and VMotion can function properly.

More information on the SAN configuration for the EMC, Cisco and VMware layouts for the environment can be found in [SAN Design Considerations for EMC RecoverPoint](#), page 37.

### VMware HA

As mentioned before, each data center had one VMware ESX cluster defined and the appropriate ESX hosts as members. VMware HA was enabled on each cluster group as the VMware Infrastructure Client output illustrates in [Figure 43](#).

**Figure 43** *VMware HA and DRS Feature Deployment*

Name  
DC1-CLS

Features

☒ Enable VMware HA

VMware HA detects failures and provides rapid recovery for the virtual machines running within a cluster. Core functionality includes host monitoring and virtual machine monitoring to minimize downtime when heartbeats are lost.

☒ Enable VMware DRS

VMware DRS enables VirtualCenter Server to manage hosts as an aggregate pool of resources. Cluster resources can be divided into smaller resource pools for users, groups, and virtual machines.

VMware DRS also enables VirtualCenter to manage the assignment of virtual machines to hosts automatically, suggesting placement when virtual machines are powered on, and migrating running virtual machines to balance load and enforce resource allocation policies.

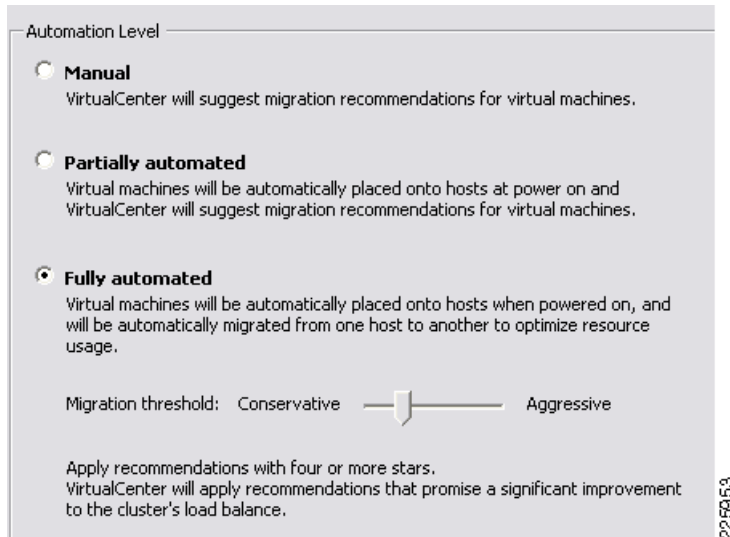
225932

The VMware HA feature is not specific to any application or OS environment such as Exchange 2007 but it offer recovery for VMs and hosts in the event of failure or scheduled maintenance by offering the ability to automatically evacuate VMs or to reboot them if their heartbeats are lost or if a host goes down unexpectedly. This adds a great deal of added availability for Exchange 2007 VMs running on the clustered hosts.

## VMware DRS

As [Figure 43](#) above illustrates, VMware DRS is enabled on the ESX clusters. DRS is a powerful tool that allows for the placement of all of the ESX host resources into a pool that is managed by Virtual Center. DRS monitors resources such as CPU and memory load on the ESX host and VMs and can simply offer recommendations on VM-to-Host placement or it can even perform automatic movement of the VM to the most appropriate host with no impact to the availability of the VM itself. As [Figure 44](#) shows, there are three settings for DRS.



**Figure 44**      **DRS Automation Setting**

In the testing of Exchange 2007 on VMware, the “Fully automated” automation level was used. The goal was to have a hands-free operation for the management of available resources. DRS works well if the VM can be relocated to any of the hosts in the cluster without any dependency on any one host. This is important because if a VM is tied to a resource that is specific to one host, such as a local disk resources, or locally attached peripheral then DRS, HA and VMotion will have issues with the automatic movement of a VM to a host that does not have the same resources the VM depends on.

Figure 45 shows an example of how DRS monitored the cluster for increases in utilization that threw the cluster resource pool out of balance. A large number of connections hit the “e2k7-cas2” VM and caused it to exceed the DRS threshold.

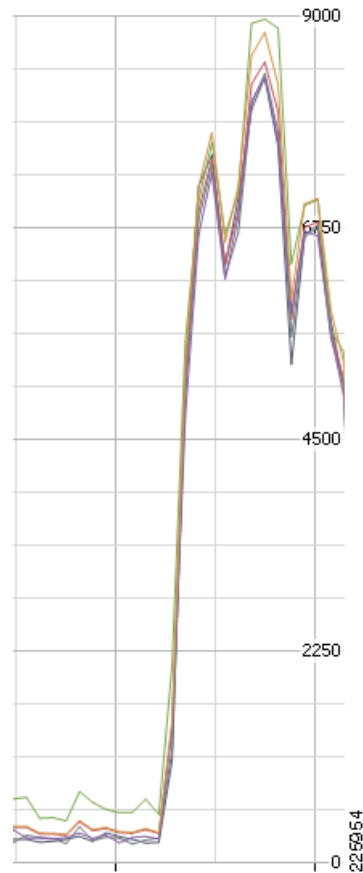

**Figure 45** Spike in CPU on ESX Host Running e2k7-cas2

Figure 46 shows that DRS moved e2k7-cas2” from the ESX host it was on due to the fact that a large number of connections hit the CAS VM and drove its CPU and memory usage above the DRS threshold. DRS automatically moved the VM to a more appropriate ESX host in the cluster.

**Figure 46** DRS Action History

| DRS Action History                                                                                                                  |                        |
|-------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| DRS Actions                                                                                                                         | Time                   |
|  Migrate E2K7-CAS2 from dcaal-esx6 to dcaal-esx5 | 12/18/2008 3:19:23 ... |

22:59:56

## VMware SRM

Site Recovery Manager provides an integrated disaster recovery workflow application that automates and controls the site to site failover of virtual machines in the event of a disaster. In testing, SRM was deployed in data center 1 (San Jose) and data center 2 (RTP) to facilitate replication and recovery of the Microsoft Exchange 2007 Mailbox Virtual Machine between the two data centers. Storage for the Exchange Server was created on three LUNs. A 45Gb operating system LUN, a 80Gb Log LUN, and a 140Gb Message Store LUN. These three LUNs were configured in a consistency group on RecoverPoint that was mirrored in the second data center by the RecoverPoint appliances. RecoverPoint consistency groups should be setup and fully initialized for the virtual machines to be protected before beginning

with the SRM install and configuration. RecoverPoint consistency group for SRM replication must not be configured for Continuous Data Replication (CDP) but only for Continuous Remote Replication (CRR). This is the only time RecoverPoint needs to be configured from its user interface.

SRM requires vCenter to be installed in both local and recovery data centers. Virtual machines within each data center must be managed by the local vCenter server. SRM is installed as a plugin to vCenter and is integrated into the vCenter interface as seen in [Figure 47](#).

Site Recovery Manager uses a Storage Recovery Adapter to interact with smart storage arrays. Each adapter supports specific manufacturer arrays. As a pre-requisite, each site requires a separate instance of either of the following databases: Microsoft SQL Server or Oracle Server. Once the database and RecoverPoint tasks are complete, installation of Site Recovery Manager can proceed.

Refer to the *H5582 VMware Site Recovery Manager with EMC RecoverPoint Implementation Guide* at the following URL for installation and configuration details:

<http://www.emc.com/collateral/software/technical-documentation/h5582-vmware-site-recovery-manager-with-recoverpoint-implguide.pdf>

## Customization

Customization with the integrated customization tools in vCenter can be used to change the IP address in the protected VM from the protected site to its new IP address in the recovered VM at the recovery site. Customizing the IP change with Site Recovery Manager helps to prevent time lost from a misconfiguration. The customization can be configured and tested before any actual recovery takes place. All groups involved with a VM can verify the configuration and the process. SRM can be configured to recover multiple VMs and having the customization scripts renumber the IP addresses of each VM greatly simplifies the recovery of multiple machines and significantly reduces the chance for human error. The net result from customizing the IP change is reduced down time and revenue. The customization setup is completed on the recovery side vCenter server.

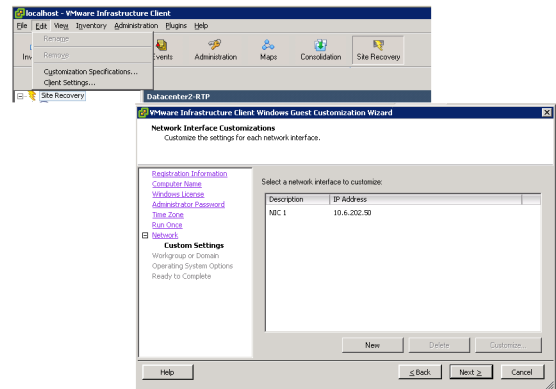


### Note

Customization is not yet supported for Windows Server 2008 as such testing was done on Windows Server 2003 for one of the MBX servers. Refer to VMware Site Recovery Manager Compatibility Matrixes at the following URL for currently supported operating system:

[http://www.vmware.com/pdf/srm\\_10\\_compat\\_matrix.pdf](http://www.vmware.com/pdf/srm_10_compat_matrix.pdf)

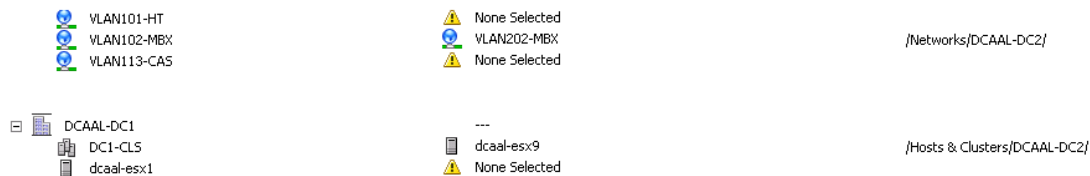
[Figure 47](#) shows the vCenter customization menus. The customized address after configuration is shown. Remember all other fields are ignored for the SRM customization and just need to be entered for the menus to progress. If failover occurs to the recovery site and failback is needed to the original datacenter, another customization script will be needed to change the IP address back at the original site.

**Figure 47 vCenter Customization Setup**

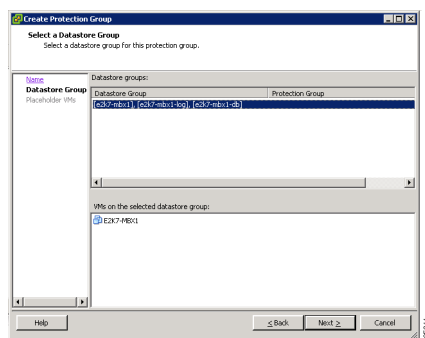
## Site Recovery Manager Configuration for Testing and Recovery

Site Recovery Manager requires only seven configuration items to setup and successfully initiate a test or full failover of a virtual machine or machines. Configuration of SRM as stated earlier is well documented by EMC, this section will highlight only the key concepts:

- 
- Step 1** Configure RecoverPoint Consistency Group for CRR, configured, initialized, and replicating between data centers.
  - Step 2** Install and configure SRM with array managers at both sites.
  - Step 3** Inventory Mappings are leveraged from the integration of existing vCenter configurations at each site into Site Recovery Manager. Manual configuration is necessary to map a resource for the recovery from one site to another. Two snippets of the Inventory Mappings screen is shown below in [Figure 48](#). In this example, the VLAN 102 MBX network, which is the primary network for the Exchange Mailbox server, maps over to the recovery network VLAN 202 MBX. Compute resources map the VM hardware at the protected site to a corresponding ESX machine at the recovery site. In this setup, the source at data center one is DC1-CLS a ESX cluster. The cluster maps over to server ESX9. Clusters do not need to map to another cluster as long as the target ESX server in the recovery data center can handle the new load of the recovered VMs.

**Figure 48** *Configure Inventory Mappings at Protected Site to Recovery Site*

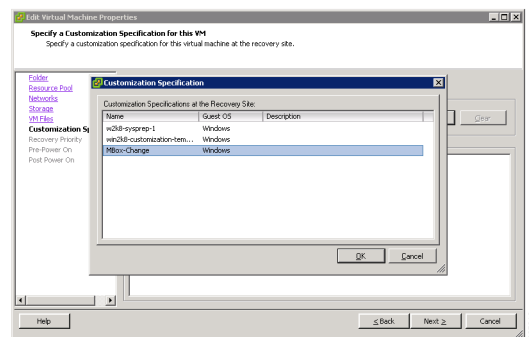
- Step 4** Creating a Protection Group on the Protected Site. Refer to [Figure 49](#). The data store seen here reflects the three LUNs that are replicated in RecoverPoint in the consistency group configured earlier for Exchange. The three LUNs e2k7-mbx1, e2k7-mbx1-log, and e2k7-mbx1-db all correspond to Exchange LUNs. Site Recovery manager in conjunction with the Site Recovery Adapter for RecoverPoint scan through the RecoverPoint configured replication groups and scan configured data stores on VMs in vCenter. When the data store group is selected, the VM associated to this group is displayed and selected. E2K7-MBX1 in this case is the correct VM. The failover example used later is E2K7-MBX2, there is no difference in the configuration of SRM for this group other than the number. Note that RecoverPoint must be fully configured with all LUNs or the Datastore will not show up in the array manager in SRM.

**Figure 49** *Create a Protection Group*

- Step 5** (Optional) Create customization script on recovery site as described earlier.

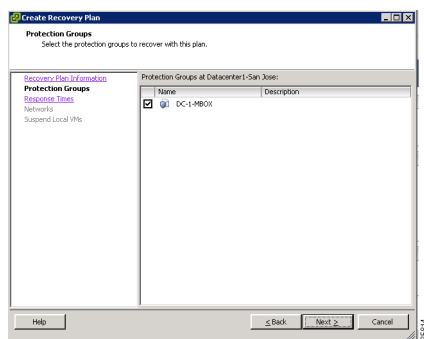
- Step 6** Configure protection for the selected VM on the Protected Site. Proceed through the dialog and verify, destination folder is selected correctly.
- When browsed the customizations that are configured on the remote vCenter server at the recovery site will be listed. Note that the customization script must exist and have been previously configured on the recovery data center vCenter server. Select the correct customization as in [Figure 50](#).
  - Priority of recovery-- All high priority VMs are recovered in sequence, all other priorities will recover in parallel. When recovering multiple VMs, it is critical to understand the SRM priorities if there are dependencies between applications as they come up. In this setup, only one VM was recovered though multiple VMs can be restored in one recovery plan.
  - There are two items that cover pre and post power on messages and custom commands that were not used during these tests.

**Figure 50** *Select Customization Specification*



- Step 7** The Recovery Plan at the Recovery site must reference the proper protection group at the Protected Site.
- In this test one VM is protected at the protected site. If there were multiple VMs protected at the protected site, they would appear here. Select the appropriate protection group from the protected site. See [Figure 51](#).

**Figure 51** *Create Recovery Plan at Recovery Site, Reference Protection Group at Protected Site*



- b. Recovery Networks and Test Networks—Testing, if left on auto will create a self-contained test bubble which will not affect any other VMs. If a network is selected here, the test will run on that network instead of a protected bubble. The effect of running a test on a production network can be problematic with servers joining domains or IP addresses showing live in load balancers when they are actually just brought up for test. If further testing is needed on the real recovery network proceed with caution.
- c. SRM has an option to suspend VMs to free resources to bring up the recovered VM. In testing, the recovery site compute resources had adequate headspace and this feature was not needed. If the recovery site was running test instances or development on the recovery resource, SRM could shutdown these low priority VMs before bringing up the recovered VM.

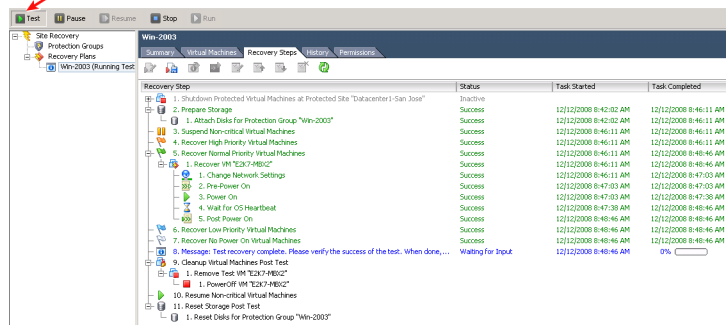
## Testing and Recovering with the "Recovery Plan"

Site Recovery Manager allows for complete testing of a recovery plan. The test can take place on a test network inside a bubble completely isolated from any other VM's that may be running on the server. The test can also take place on a specified network if the VM needs to reach other servers. The test features of site recovery manager allow for a full and thorough test from IP customization to application turn-up complete with the replicated storage. Storage is brought to the test with the Storage Recovery Adapter interface to RecoverPoint. RecoverPoint make the replicated image available to the recovery VM. Any writes to this image are lost upon test completion. Each step of the process is documented for compliance and a history of test success or failure is easily available from within vCenter with the SRM plug-in. Test the recovery plan by selecting the green test button on the recovery site SRM. The test will run a complete test without shutting down the VMs at the protected site. When the VM comes up, the script will pause for user verification. At this point, storage can be verified and applications fully tested.

Figure 52 shows output from running the **test** command.

**Figure 52**      **Testing Recovery Plan**

Press The Green  
Test Button



To execute a full recovery of a VM refer to [Figure 52](#) above and select the run button. Note that certain steps are executed for test only or for recovery only. In a full recovery, VMs at the protected site are shutdown before being recovered at their recovery site. The same steps for customization and bringing the VM up and running are the same as in test. The main difference is with RecoverPoint. SRM will initiate a production failover to the replicated site. This storage, unlike in test, is now fully writable and active as the primary storage. If the previously protected site is still up, RecoverPoint will start journaling data the opposite direction. [Figure 53](#) shows the RecoverPoint GUI status page before failover. Note that SJ Data Center 1 is in production.



**Figure 53** *RecoverPoint GUI before SRM Recovery*

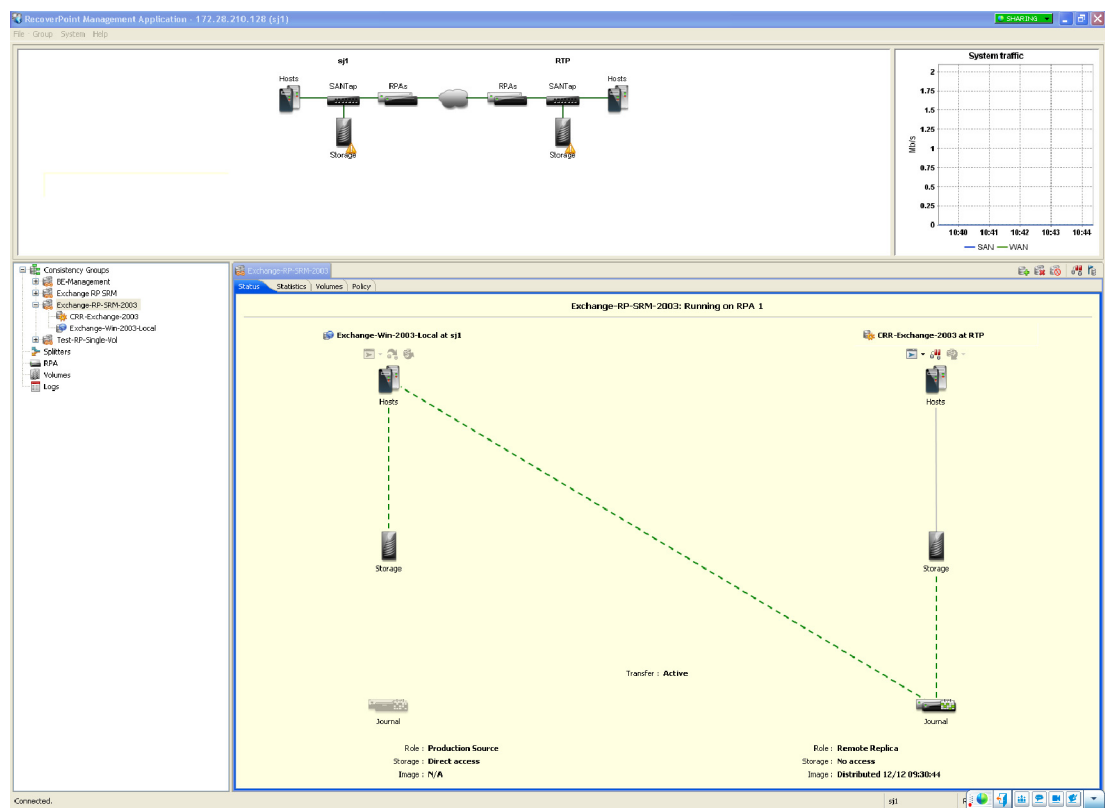


Figure 54 shows the results for the same recovery plan, but in non-test full failover.

**Figure 54 SRM Recovery****Win-2003**

VMware Site Recovery Manager

**Description**

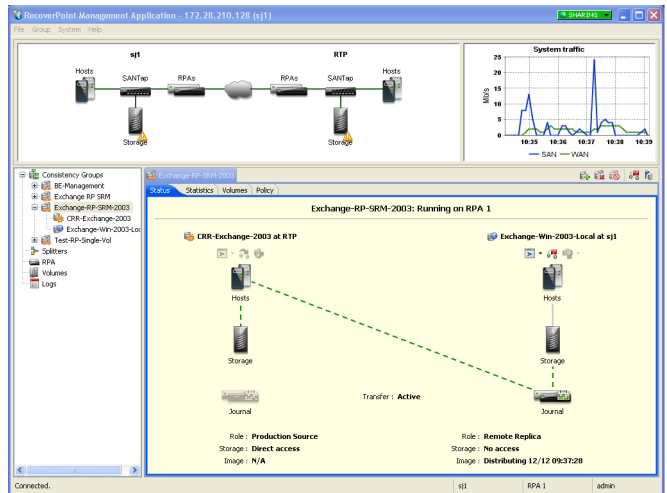
Start Time: 12/12/2008 9:16:26 AM  
 Finish Time: 12/12/2008 9:30:40 AM  
 Total Execution Time: 00:14:13  
 Mode: Recovery  
 Overall Result: Success

| Recovery Step                                                                                                                                          | Result  | Execution Time |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------|----------------|
| 1. Shutdown Protected Virtual Machines at Protected Site "Datacenter1-San Jose"                                                                        | Success | 00:00:21       |
| 1.1. Shutdown Low Priority Protected Virtual Machines                                                                                                  | Success | 00:00:00       |
| 1.2. Shutdown Normal Priority Protected Virtual Machines                                                                                               | Success | 00:00:21       |
| 1.2.1. Shutdown Protected Site VM "E2K7-MBX2"                                                                                                          | Success | 00:00:21       |
| 1.2.1.1. Shutdown Guest OS for Remote VM "E2K7-MBX2"                                                                                                   | Success | 00:00:01       |
| 1.2.1.2. Wait for Guest OS Shutdown                                                                                                                    | Success | 00:00:20       |
| 1.2.1.3. Power off VM "E2K7-MBX2"                                                                                                                      | Success | 00:00:00       |
| 1.3. Shutdown High Priority Protected Virtual Machines                                                                                                 | Success | 00:00:00       |
| 2. Prepare Storage                                                                                                                                     | Success | 00:04:05       |
| 2.1. Attach Disks for Protection Group "Win-2003"                                                                                                      | Success | 00:04:05       |
| 3. Suspend Non-critical Virtual Machines                                                                                                               | Success | 00:00:00       |
| 4. Recover High Priority Virtual Machines                                                                                                              | Success | 00:00:00       |
| 5. Recover Normal Priority Virtual Machines                                                                                                            | Success | 00:09:45       |
| 5.1. Recover VM "E2K7-MBX2"                                                                                                                            | Success | 00:09:45       |
| 5.1.1. Change Network Settings                                                                                                                         | Success | 00:08:55       |
| 5.1.2. Pre-Power On                                                                                                                                    | Success | 00:00:00       |
| 5.1.3. Power On                                                                                                                                        | Success | 00:00:01       |
| 5.1.4. Wait for OS Heartbeat                                                                                                                           | Success | 00:00:48       |
| 5.1.5. Post Power On                                                                                                                                   | Success | 00:00:00       |
| 6. Recover Low Priority Virtual Machines                                                                                                               | Success | 00:00:00       |
| 7. Recover No Power On Virtual Machines                                                                                                                | Success | 00:00:00       |
| 8. Message: Test recovery complete. Please verify the success of the test. When done, click Continue to clean up the test and return to a ready state. |         |                |
| 9. Cleanup Virtual Machines Post Test                                                                                                                  |         |                |
| 9.1. Remove Test VM "E2K7-MBX2"                                                                                                                        |         |                |
| 9.1.1. PowerOff VM "E2K7-MBX2"                                                                                                                         |         |                |
| 10. Resume Non-critical Virtual Machines                                                                                                               |         |                |
| 11. Reset Storage Post Test                                                                                                                            |         |                |
| 11.1. Reset Disks for Protection Group "Win-2003"                                                                                                      |         |                |

225821

Actual recovery with SRM yields a faster configuration change than observed in testing. A screen shot of RecoverPoint in [Figure 55](#) now shows the RTP Data Center 2 site as the active production site and Data Center 1 SJ as the recovery site. SRM has made RTP the production source without any intervention to RecoverPoint outside of SRM.

Exchange came up and was running after the full failover. Mail that was sent before the failover was on the recovered VM. All mail was able to be sent and received from the new recovered server. SRM completed the recover with a click of a button while providing a full audit trail and not needing to bring server admins, network admins, or storage admins.

**Figure 55 RecoverPoint GUI after SRM Recovery**

Failback is not available in SRM Version 1.0 Build 1. To failback to data center 1 from data center 2, complete the following steps:

- Step 1** Remove recovery plans from both SRM instances.
- Step 2** Remove protection group at original protected site.
- Step 3** Unregister the original protected VM at the protected site.
- Step 4** Update array manager configuration at the new protected side.
- Step 5** Add protection group for the new recovered VM and configure protection for VM.
- Step 6** Configure customization as desired at new recovery site.
- Step 7** Create recovery plan at new protected site.
- Step 8** Select recovery plan at new recovery site.
- Step 9** Test failover.
- Step 10** Execute failover.
- Step 11** Repeat steps 1 to 9 again once the VM is back at the original protected site

**Note**

SRM removes all log files corresponding to the recovery plan when the plan is deleted. Export your logs before removing your recovery plans.

## Considerations and Results Observed During Setup and Testing

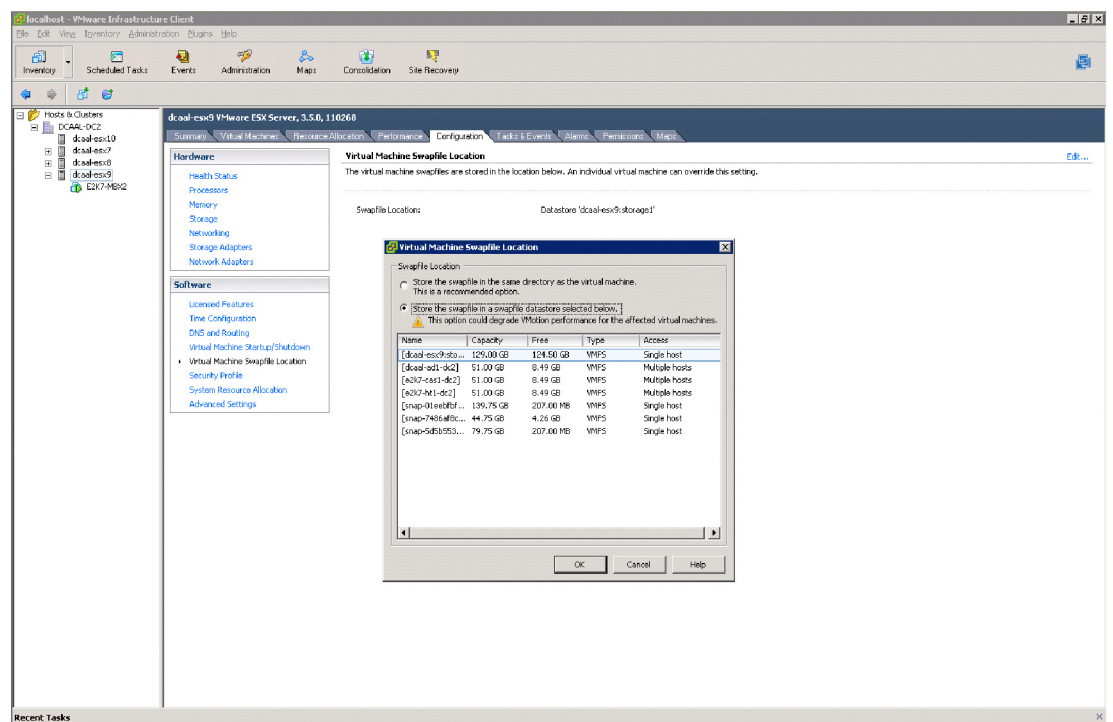
### Swap space by default uses same folder as vmdk boot file

This caused an issue as the replicated LUN that held the **vmdk** file was not large enough to handle the additional swap space

**Solution** The following two optional solutions; Option 1 was used in this test scenario.

1. Option 1. Use local storage on ESX server for swap space for VM. In vCenter, you can configure the swap file location under **Configuration > Virtual Machine Swap File Location** as shown in .
2. Option 2. Manually create memory reservations on the placeholder VM after creating the protection group. This ensures correct memory allocations and prevents the need for extra storage space on the VMFS volume storing the **vmdk** boot file.

**Figure 56** *Edit Swap File Location for Workaround*

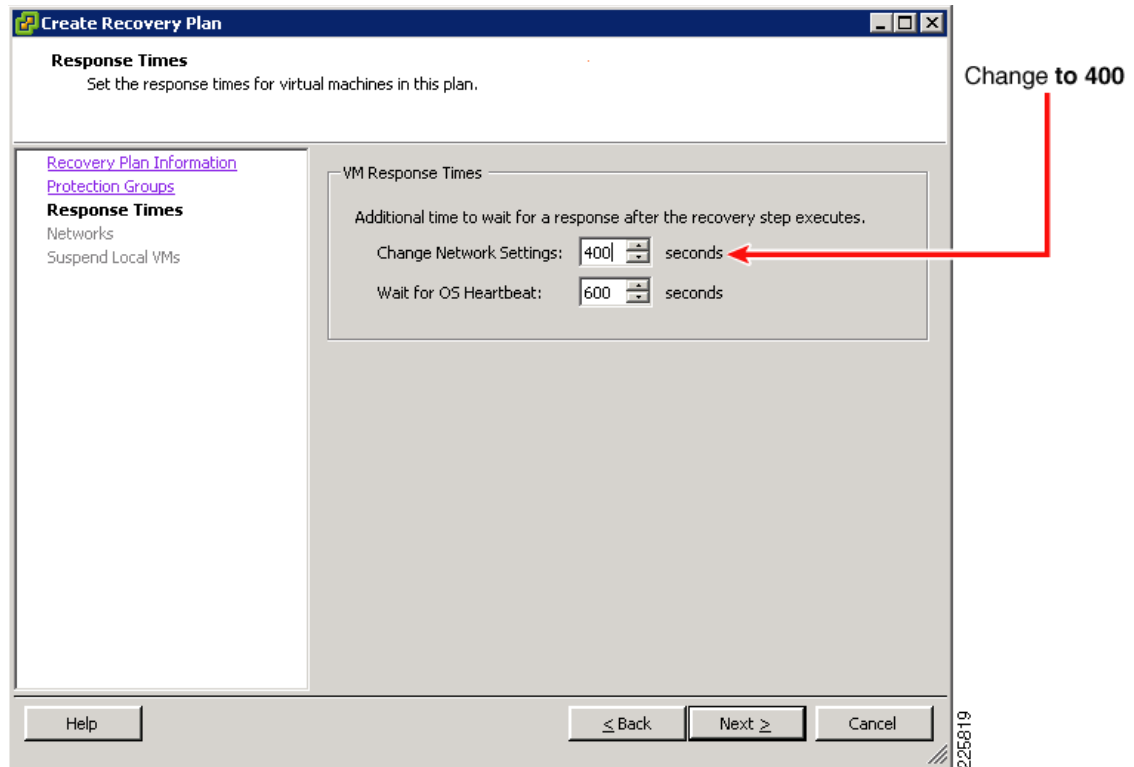


### SRM had timing issues with VMWare tools coming up during reconfiguration of the custom IP address.

There are known timing issues that exist with reconfiguration of the virtual machine in Site Recovery Manager. Some fixes for this problem should reduce this time delay are to be release in the upcoming Patch 11 for ESX. To prevent the recovery plan from failing the fix was to extend the Change Network Settings to 400 seconds. Refer to [Figure 57](#) to see how the **Recovery Plan Timers** are updated. This change allows everything to timeout properly during the test and allow the recovery plan to report success. If this change is not made, the recovery plan will fail even though the changes were made properly.

Note that high priority VMs are executed serially in the recovery plan. This delay on IP change could severely impact testing even a small number of high priority VMs in a recovery plan.

**Figure 57** *Change Recovery Plan Response Times*



SCSI reservation conflicts were significantly slowing down testing and marginally slowing down recovery tests.

Site Recovery manager performs a Rescan HBA Process each time it does a recovery or recovery test, in addition if it is a test, it goes through it again when it resets the storage. The problem arises from RecoverPoint owning the access to the LUNs until SRM activates a copy for ESX to use. As of the date when this testing was completed, when a rescan is issued, if there are SCSI reservations, the rescan can take a long time because the SCSI commands are retried 80 times before they timeout. This is the normal scenario how it works with regards to ESX. VMware is aware of this issue and this feature will be changed in future ESX releases. Fortunately, the reset of the storage only occurs during testing and adds about 10 to 12 minutes to the test. During an actual failover the delay at the start is about 2 minutes. Disk cleanup delays are not observed in full recovery because it only executes in test mode; there is no known workaround. Also note that high priority VMs are executed serially in the recovery plan so this delay will compound the delay as each VM is brought up.

Figure 58 show results of a failover test with IP customization.

**Figure 58 Recovery Test with Customization**

Win-2003 VMware Site Recovery Manager

Description

Start Time: 12/12/2008 7:50:55 AM  
 Finish Time: 12/12/2008 8:32:53 AM  
 Total Execution Time: 00:41:57  
 Mode: Test  
 Overall Result: Success

Note this timer is to change IP.

| Recovery Step                                                                                                                                          | Result  | Execution Time |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------|----------------|
| 1. Shutdown Protected Virtual Machines at Protected Site "Datacenter1-San Jose"                                                                        |         |                |
| 1.1. Shutdown Low Priority Protected Virtual Machines                                                                                                  |         |                |
| 1.2. Shutdown Normal Priority Protected Virtual Machines                                                                                               |         |                |
| 1.2.1. Shutdown Protected Site VM "E2K7-MBX2"                                                                                                          |         |                |
| 1.2.1.1. Shutdown Guest OS for Remote VM "E2K7-MBX2"                                                                                                   |         |                |
| 1.2.1.2. Wait for Guest OS Shutdown                                                                                                                    |         |                |
| 1.2.1.3. Power off VM "E2K7-MBX2"                                                                                                                      |         |                |
| 1.3. Shutdown High Priority Protected Virtual Machines                                                                                                 |         |                |
| 2. Prepare Storage                                                                                                                                     | Success | 00:04:20       |
| 2.1. Attach Disks for Protection Group "Win-2003"                                                                                                      | Success | 00:04:19       |
| 3. Suspend Non-critical Virtual Machines                                                                                                               | Success | 00:00:00       |
| 4. Recover High Priority Virtual Machines                                                                                                              | Success | 00:00:00       |
| 5. Recover Normal Priority Virtual Machines                                                                                                            | Success | 00:10:58       |
| 5.1. Recover VM "E2K7-MBX2"                                                                                                                            | Success | 00:10:58       |
| 5.1.1. Change Network Settings                                                                                                                         | Success | 00:09:37       |
| 5.1.2. Pre-Power On                                                                                                                                    | Success | 00:00:00       |
| 5.1.3. Power On                                                                                                                                        | Success | 00:00:04       |
| 5.1.4. Wait for OS Heartbeat                                                                                                                           | Success | 00:01:16       |
| 5.1.5. Post Power On                                                                                                                                   | Success | 00:00:00       |
| 6. Recover Low Priority Virtual Machines                                                                                                               | Success | 00:00:00       |
| 7. Recover No Power On Virtual Machines                                                                                                                | Success | 00:00:00       |
| 8. Message: Test recovery complete. Please verify the success of the test. When done, click Continue to clean up the test and return to a ready state. | Success | 00:09:51       |
| 9. Cleanup Virtual Machines Post Test                                                                                                                  | Success | 00:00:04       |
| 9.1. Remove Test VM "E2K7-MBX2"                                                                                                                        | Success | 00:00:04       |
| 9.1.1. PowerOff VM "E2K7-MBX2"                                                                                                                         | Success | 00:00:04       |
| 10. Resume Non-critical Virtual Machines                                                                                                               | Success | 00:00:00       |
| 11. Revert Storage Post Test                                                                                                                           | Success | 00:16:43       |
| 11.1. Reset Disks for Protection Group "Win-2003"                                                                                                      | Success | 00:16:43       |

2256818

Disk Cleanup, SCSI conflicts

## Summary

The Cisco, EMC, and VMware virtual solution provides for a scalable Microsoft Exchange Server 2007 environment brought together by the industry leaders in server virtualization, networking, and data storage that address common challenges facing IT professionals. By virtualizing the Exchange Server 2007 environment using the VMware Infrastructure 3 (VI3) platform and EMC storage arrays with RecoverPoint replication appliances, organizations can simplify disaster recovery and build a more flexible and reliable infrastructure. This reduces the overall complexity and makes email messaging increasingly more available. The solution also simplifies management, diminishes a variety of IT expenditures, helps IT professionals take full advantage of existing IT assets, and improves performance of an Exchange deployment as a whole. Cisco, EMC, and VMware collaborated to validate this solution to provide design and deployment guidance for those customers looking to fully virtualize Exchange 2007.

# Appendix

## Reference Documents

- Microsoft Exchange Server 2007:  
<http://www.microsoft.com/exchange/default.mspx>
- EMC Solutions for Microsoft Exchange:  
<http://www.emc.com/solutions/application-environment/microsoft/solutions-for-microsoft-exchange-unified-communications.htm>
- Cisco Connection Online—Data Center:  
<http://www.cisco.com/go/dc>
- Cisco Validated Design (CVD) Zone:  
<http://www.cisco.com/go/designzone>
- Integrating Microsoft Exchange Server in a Cisco multisite data center design:  
<http://www.cisco.com/en/US/docs/solutions/Verticals/mstdcmsftex.html>
- Cisco IronPort secure email appliances:  
[http://www.ironport.com/products/email\\_security\\_appliances.html](http://www.ironport.com/products/email_security_appliances.html)
- Cisco SAN and MDS design:  
[http://www.cisco.com/en/US/products/ps5990/prod\\_white\\_papers\\_list.html](http://www.cisco.com/en/US/products/ps5990/prod_white_papers_list.html)
- Cisco application performance and site-selection solutions:  
[http://www.cisco.com/en/US/products/ps5719/Products\\_Sub\\_Category\\_Home.html](http://www.cisco.com/en/US/products/ps5719/Products_Sub_Category_Home.html)  
<http://www.cisco.com/en/US/products/hw/contnetw/ps4162/index.html>
- Cisco ASA security solutions:  
[http://www.cisco.com/en/US/products/ps6120/prod\\_brochure\\_list.html](http://www.cisco.com/en/US/products/ps6120/prod_brochure_list.html)
- SAN system design and deployment guide:  
<http://www.vmware.com/resources/techresources/772>
- Performance tuning best practices for ESX Server 3:  
<http://www.vmware.com/resources/techresources/707>
- CLARiiON integration with VMware ESX Server:  
<http://www.vmware.com/resources/techresources/241>
- VMware certified compatibility guides:  
<http://www.vmware.com/resources/guides.html>
- Deploying Exchange on a VMware platform:  
[http://www.vmware.com/landing\\_pages/exchange\\_resources.html](http://www.vmware.com/landing_pages/exchange_resources.html)
- Virtual solution for Microsoft Exchange Server 2007 using VMware Infrastructure 3 and EMC CLARiiON CX3-20 Small Computer System Interface over IP (iSCSI) storage:  
[http://www.vmware.com/files/pdf/exchange\\_solution\\_overview.pdf](http://www.vmware.com/files/pdf/exchange_solution_overview.pdf)
- VMware Site Recovery Manager with EMC Recover Point Implementation Guide

<http://www.emc.com/collateral/software/technical-documentation/h5582-vmware-site-recovery-manager-with-recoverpoint-implguide.pdf>

- VMware Site Recovery Manager Compatibility Matrixes

[http://www.vmware.com/pdf/srm\\_10\\_compat\\_matrix.pdf](http://www.vmware.com/pdf/srm_10_compat_matrix.pdf)

- Cisco MDS 9000 Family Storage Services Module Software Installation and Upgrade Guide

[http://www.cisco.com/en/US/docs/storage/san\\_switches/mds9000/sw/ssm/upgrade/guide/SSMupgd.html](http://www.cisco.com/en/US/docs/storage/san_switches/mds9000/sw/ssm/upgrade/guide/SSMupgd.html)

EMC PowerLink <http://powerlink.emc.com>

## Complete ACE Configuration

```
crypto csr-params testparams
 country US
 state California
 locality SJ
 organization-name AS
 organization-unit TAS
 common-name www.testssl.com
 serial-number cisco123
access-list BPDU-Allow ethertype permit bpd

access-list ANYONE line 10 extended permit ip any any
access-list ANYONE line 20 extended permit icmp any any

kalap udp
 ip address 10.5.25.10 encryption md5 ese
 ip address 10.6.25.10 encryption md5 ese

probe icmp PING
 interval 5
 passdetect interval 2
 passdetect count 1

parameter-map type http TCP-REUSE
 server-conn reuse
 persistence-rebalance
parameter-map type connection conn-map
 tcp-options selective-ack allow
 tcp-options timestamp allow
 tcp-options window-scale allow

action-list type modify http sslrewrite
 ssl url rewrite location ".*" sslport 443 clearport 80

rserver host CAS1
 description CAS Server 1
 ip address 10.5.103.50
 inservice
rserver host CAS2
 description CAS Server 2
 ip address 10.5.103.51
 inservice
rserver redirect REDIRECT-TO-HTTPS
 webhost-redirection https://owa.ese.com/owa 302
```



```

inservice

ssl-proxy service SSL-OFFLOAD
 key host.key
 cert host.cert

serverfarm host OWA
 probe PING
 rserver CAS1 80
 inservice
 rserver CAS2 80
 inservice
serverfarm redirect REDIRECT-SERVERFARM
 rserver REDIRECT-TO-HTTPS
 inservice

sticky http-cookie ACE-Insert sticky-cookie-group
 cookie insert
 timeout 20
 replicate sticky
 serverfarm OWA
sticky ip-netmask 255.255.255.255 address source IP-Sticky
 replicate sticky
 serverfarm OWA

class-map type management match-any KALAP-MGMT
 2 match protocol kalap-udp source-address 10.5.25.10 255.255.255.255
 3 match protocol kalap-udp source-address 10.6.25.10 255.255.255.255

class-map match-all OWA-VIP
 2 match virtual-address 10.5.103.11 tcp eq www
class-map match-all OWA-VIP-443
 2 match virtual-address 10.5.103.11 tcp eq https
class-map type management match-any REMOTE-MGT
 201 match protocol snmp any
 202 match protocol http any
 203 match protocol https any
 204 match protocol icmp any
 205 match protocol ssh any
 206 match protocol telnet any

policy-map type management first-match KALAP-MGMT-POL
 class KALAP-MGMT
 permit
policy-map type management first-match MANAGEMENT
 class REMOTE-MGT
 permit

policy-map type loadbalance first-match OWA-LB-POLICY
 class class-default
 serverfarm OWA
policy-map type loadbalance first-match REDIRECT-PM
 class class-default
 serverfarm REDIRECT-SERVERFARM
policy-map type loadbalance http first-match STICKYLB
 class class-default
 sticky-serverfarm sticky-cookie-group
 action sslrewrite
policy-map type loadbalance first-match STICKY_IP_LB
 class class-default
 sticky-serverfarm IP-Sticky
 action sslrewrite

policy-map multi-match OWA-POLICY-MAP

```

```

class OWA-VIP-443
 loadbalance vip inservice
 loadbalance policy STICKY_IP_LB
 loadbalance vip icmp-reply active
 loadbalance vip advertise active
 appl-parameter http advanced-options TCP-REUSE
 ssl-proxy server SSL-OFFLOAD
 connection advanced-options conn-map
class OWA-VIP
 loadbalance vip inservice
 loadbalance policy REDIRECT-PM
 loadbalance vip icmp-reply active
 loadbalance vip advertise active
 appl-parameter http advanced-options TCP-REUSE
 connection advanced-options conn-map

interface vlan 82
 description Mgmt interface
 ip address 172.28.210.99 255.255.255.0
 no icmp-guard
 access-group input ANYONE
 service-policy input MANAGEMENT
 no shutdown
interface vlan 103
 description OWA-Client-side-vlan
 bridge-group 1
 access-group input BPDU-Allow
 access-group input ANYONE
 service-policy input MANAGEMENT
 service-policy input OWA-POLICY-MAP
 service-policy input KALAP-MGMT-POL
 no shutdown
interface vlan 113
 description Server-side-vlan
 bridge-group 1
 access-group input BPDU-Allow
 access-group input ANYONE
 no shutdown

interface bvi 1
 ip address 10.5.103.4 255.255.255.0
 description OWA-Bridged-vlans
 no shutdown

ip route 0.0.0.0 0.0.0.0 10.5.103.1
ip route 172.0.0.0 255.0.0.0 172.28.210.1

```

