



Newer Cisco SBA Guides Available

This guide is part of an older series of Cisco Smart Business Architecture designs. To access the latest Cisco SBA Guides, go to <http://www.cisco.com/go/sba>

Cisco strives to update and enhance SBA guides on a regular basis. As we develop a new series of SBA guides, we test them together, as a complete system. To ensure the mutual compatibility of designs in Cisco SBA guides, you should use guides that belong to the same series.





SBA

DATA CENTER

DEPLOYMENT
GUIDE

Advanced Server-Load Balancing Deployment Guide

● ● ● SMART BUSINESS ARCHITECTURE

August 2012 Series

Preface

Who Should Read This Guide

This Cisco® Smart Business Architecture (SBA) guide is for people who fill a variety of roles:

- Systems engineers who need standard procedures for implementing solutions
- Project managers who create statements of work for Cisco SBA implementations
- Sales partners who sell new technology or who create implementation documentation
- Trainers who need material for classroom instruction or on-the-job training

In general, you can also use Cisco SBA guides to improve consistency among engineers and deployments, as well as to improve scoping and costing of deployment jobs.

Release Series

Cisco strives to update and enhance SBA guides on a regular basis. As we develop a series of SBA guides, we test them together, as a complete system. To ensure the mutual compatibility of designs in Cisco SBA guides, you should use guides that belong to the same series.

The Release Notes for a series provides a summary of additions and changes made in the series.

All Cisco SBA guides include the series name on the cover and at the bottom left of each page. We name the series for the month and year that we release them, as follows:

month year Series

For example, the series of guides that we released in August 2012 are the “August 2012 Series”.

You can find the most recent series of SBA guides at the following sites:

Customer access: <http://www.cisco.com/go/sba>

Partner access: <http://www.cisco.com/go/sbachannel>

How to Read Commands

Many Cisco SBA guides provide specific details about how to configure Cisco network devices that run Cisco IOS, Cisco NX-OS, or other operating systems that you configure at a command-line interface (CLI). This section describes the conventions used to specify commands that you must enter.

Commands to enter at a CLI appear as follows:

```
configure terminal
```

Commands that specify a value for a variable appear as follows:

```
ntp server 10.10.48.17
```

Commands with variables that you must define appear as follows:

```
class-map [highest class name]
```

Commands shown in an interactive example, such as a script or when the command prompt is included, appear as follows:

```
Router# enable
```

Long commands that line wrap are underlined. Enter them as one command:

```
wrr-queue random-detect max-threshold 1 100 100 100 100 100  
100 100 100
```

Noteworthy parts of system output or device configuration files appear highlighted, as follows:

```
interface Vlan64  
ip address 10.5.204.5 255.255.255.0
```

Comments and Questions

If you would like to comment on a guide or ask questions, please use the [SBA feedback form](#).

If you would like to be notified when new comments are posted, an RSS feed is available from the SBA customer and partner pages.

Table of Contents

What's In This SBA Guide.....	1	Deployment Details.....	4
Cisco SBA Data Center	1	Configuring Cisco ACE Connectivity to Switches	6
Route to Success	1	Configuring Cisco ACE Appliance Network	8
About This Guide	1	Setting up Load Balancing for HTTP Servers	11
Introduction.....	2	Load-Balancing and SSL-Offloading for HTTPS Servers.....	17
Business Overview.....	2	Appendix A: Product List	22
Technology Overview.....	2	Appendix B: Configuration	23
		Appendix C: Changes.....	26

What's In This SBA Guide

Cisco SBA Data Center

Cisco SBA helps you design and quickly deploy a full-service business network. A Cisco SBA deployment is prescriptive, out-of-the-box, scalable, and flexible.

Cisco SBA incorporates LAN, WAN, wireless, security, data center, application optimization, and unified communication technologies—tested together as a complete system. This component-level approach simplifies system integration of multiple technologies, allowing you to select solutions that solve your organization's problems—without worrying about the technical complexity.

Cisco SBA Data Center is a comprehensive design that scales from a server room to a data center for networks with up to 10,000 connected users. This design incorporates compute resources, security, application resiliency, and virtualization.

Route to Success

To ensure your success when implementing the designs in this guide, you should first read any guides that this guide depends upon—shown to the left of this guide on the route below. As you read this guide, specific prerequisites are cited where they are applicable.

About This Guide

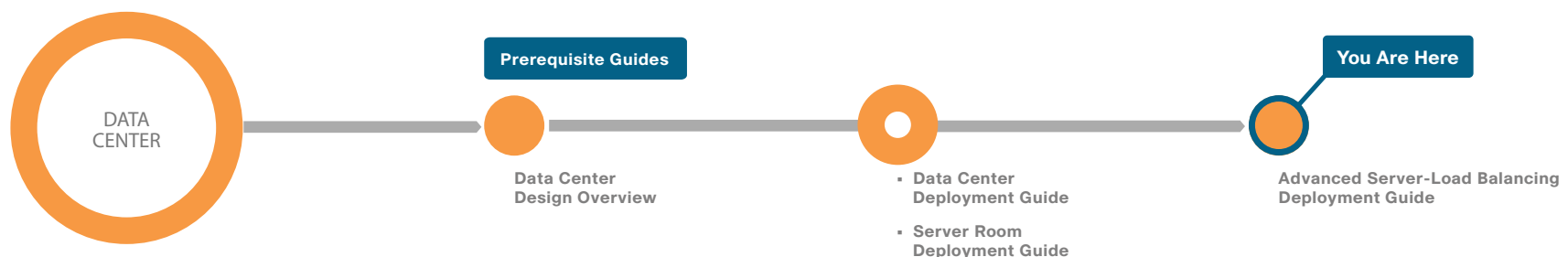
This *deployment guide* contains one or more deployment chapters, which each include the following sections:

- **Business Overview**—Describes the business use case for the design. Business decision makers may find this section especially useful.
- **Technology Overview**—Describes the technical design for the business use case, including an introduction to the Cisco products that make up the design. Technical decision makers can use this section to understand how the design works.
- **Deployment Details**—Provides step-by-step instructions for deploying and configuring the design. Systems engineers can use this section to get the design up and running quickly and reliably.

You can find the most recent series of Cisco SBA guides at the following sites:

Customer access: <http://www.cisco.com/go/sba>

Partner access: <http://www.cisco.com/go/sbachannel>



Introduction

Business Overview

The network is playing an increasingly important role in the success of a business. Key applications, such as enterprise resource planning, e-commerce, email, and portals, must be available around-the-clock to provide uninterrupted business services. However, the availability of these applications is often threatened by network overloads, as well as server and application failures. Furthermore, resource utilization is often out of balance, resulting in the low-performance resources being overloaded with requests while the high-performance resources remain idle. Application performance, as well as availability, directly affects employee productivity and the bottom-line of a company. As more users work more hours while using key business applications, it becomes even more important to address application availability and performance issues to ensure achievement of business processes and objectives.

Some of the factors that make applications difficult to deploy and deliver effectively over the network include:

- **Inflexible application infrastructure**—Application infrastructure design has historically been done on an application-by-application basis. This means that the infrastructure used for a particular application is often unique to that application. This type of design tightly couples the application to the infrastructure and offers little flexibility. Because the application and infrastructure are tightly coupled, it is difficult to partition resources and levels of control to match changing business requirements.
- **Server availability and load**—The mission-critical nature of applications puts a premium on server availability. Despite the benefits of server virtualization technology, the number of physical servers continues to grow based on new application deployments, which in turn increases power and cooling requirements.
- **Application security and compliance**—Many of the new threats to network security are the result of application- and document-embedded attacks that compromise application performance and availability. Such attacks can also potentially cause the loss of vital application data, while leaving networks and servers unaffected.

One possible solution to improve application performance and availability is to rewrite the application completely to make it network-optimized. However, this requires application developers to have a deep understanding of how different applications respond to bandwidth constraints, delay, jitter, and other network variances. In addition, developers need to accurately predict each end-user's foreseeable access method. This is simply not feasible for every business application, particularly traditional applications that took years to write and customize.

Technology Overview

The idea of improving application performance began in the data center. The Internet boom ushered in the era of the server load balancers (SLBs). SLBs balance the load on groups of servers to improve server response to client requests, and have evolved to take on additional responsibilities, such as application proxies and complete Layer 4 through 7 application switching.

Cisco Application Control Engine (Cisco ACE) is the latest SLB offering from Cisco. Its main role is to provide Layer 4 through 7 switching, but Cisco ACE also provides an array of acceleration and server offload benefits, including TCP-processing offload, SSL-processing offload, and compression. The Cisco ACE appliance sits in the data center in front of the application servers and provides a range of services to maximize server and application availability, security, and asymmetric application acceleration (from server to client browser). As a result, Cisco ACE gives IT departments more control over application and server infrastructure, which enables them to manage and secure application services more easily and improve performance.

Cisco ACE provides the following benefits:

- **Scalability**—Cisco ACE scales the performance of a server-based program, such as a web server, by distributing its client requests across multiple servers, known as a server farm. As traffic increases, additional servers can be added to the farm. With the advent of server virtualization, application servers can be staged and added dynamically as capacity requirements change.

- **High availability**—Cisco ACE provides high availability by automatically detecting the failure of a server and repartitioning client traffic among the remaining servers within seconds, while providing users with continuous service.
- **Application acceleration**—Cisco ACE improves application performance and reduces response time by minimizing latency and compressing data transfers for any HTTP-based application, for any internal or external end-user.
- **Server offload**—Cisco ACE offloads TCP processing, SSL processing, and compression from the server, which allows the server to handle more requests so more users can be served, and reduces bandwidth requirements by up to 90% without increasing the number of servers. Running SSL on the web application servers is a tremendous drain on server resources. By offloading SSL processing, those resources can be applied to traditional web-application functions. In addition, because persistence information used by the content switches is inside the HTTP header, this information is no longer visible when carried inside SSL sessions. By terminating these sessions before applying content switching decisions, persistence options become available for secure sites.
- **Flexible licensing model**—Cisco ACE is available in a number of performance options, from 500 Mbps to 4 Gbps of throughput, depending on which license is purchased. You can purchase a 1 Gbps license for your Cisco ACE appliance and then, as your performance requirements increase, upgrade the same hardware to 4 Gbps with a new license.
- **Health monitoring**—Cisco ACE uses both active and passive techniques to monitor server health. By periodically probing servers and monitoring the return traffic from the real servers, Cisco ACE rapidly detects server failures and quickly reroutes connections to available servers. A variety of health-checking features are supported, including the ability to verify web servers, SSL servers, application servers, databases, FTP servers, and streaming media servers.
- **Effective content allocation**—Cisco ACE may be used to push requests for cacheable content, such as image files, to a set of caches that can serve them more cost-effectively than the application servers.

Cisco ACE can be used to partition components of a single web application across several application server clusters. For example, the URLs, `www.mycompany.com/quotes/getquote.jsp` and `www.mycompany.com/trades/order.jsp`, could be located on two different server clusters even though the domain name is the same. This partitioning allows the application developer to easily scale the application to several servers without numerous code modifications. Furthermore, it maximizes the cache coherency of the servers by keeping requests for the same pages on the same servers.

There are several ways to integrate Cisco ACE into the data center network. Logically, the Cisco ACE appliance is deployed in front of the application cluster. Requests to the application cluster are directed to a virtual IP address (VIP) configured on the appliance. Cisco ACE receives connections and HTTP-requests, and routes them to the appropriate application server based on configured policies.

Physically, the network topology can take many forms. One-armed mode is the simplest deployment method, in which the Cisco ACE is connected off to the side of the layer 2/layer 3 infrastructure. It is not directly in the path of traffic flow and receives only traffic that is specifically intended for it. Traffic, which should be directed to it, is controlled by careful design of VLANs, virtual server addresses, server default gateway selection, or policy routes on the layer 2/layer 3 switch.

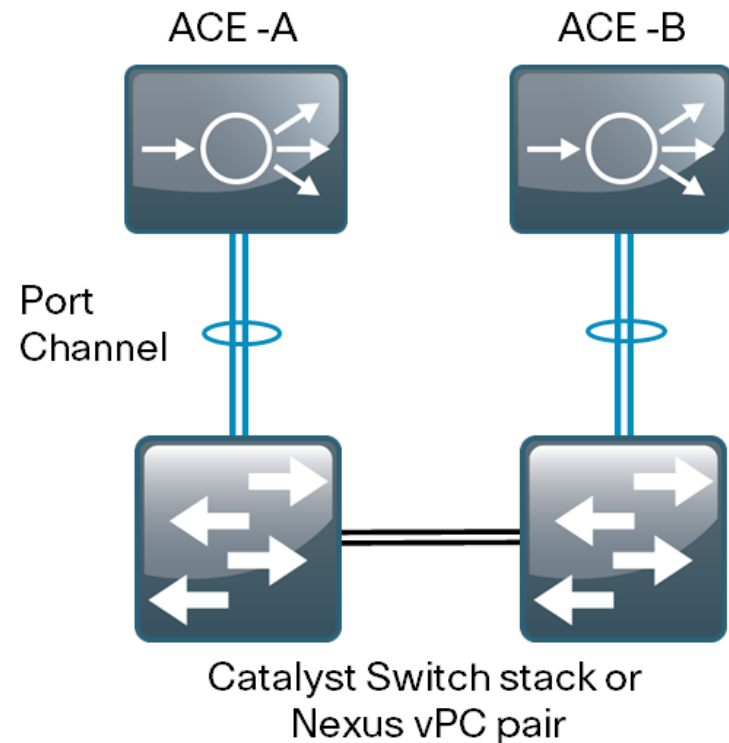
Deployment Details

Cisco ACE 4710 hardware is always deployed in pairs for highest availability, with one primary and one secondary appliance. If the primary Cisco ACE appliance fails, the secondary appliance takes control. Depending on how session-state redundancy is configured, this failover may take place without disrupting the client-to-server connection.

Cisco ACE can connect to the Cisco SBA architecture in several places, such as the data center, server room, or Internet edge DMZ, to provide application and server load-balancing services. This guide outlines two connectivity options that cover the majority of deployment scenarios.

As illustrated in Figure 1, Cisco ACE 4710 appliances are deployed in a pair for high availability. Each appliance has a port channel that is connected to the switch to scale performance. In Figure 1, the appliance uses two links for 2Gbps of available throughput, but two additional gigabit ports are available. By using four ports, the Cisco ACE appliance can scale the solution to 4Gbps. Cisco ACE operates in an active standby mode, and to maintain performance in a failure scenario, all of the links from each Cisco ACE appliance connect to only a single switch. This prevents the scenario in which Cisco ACE is connected to both switches, and a switch failure cuts the available bandwidth in half.

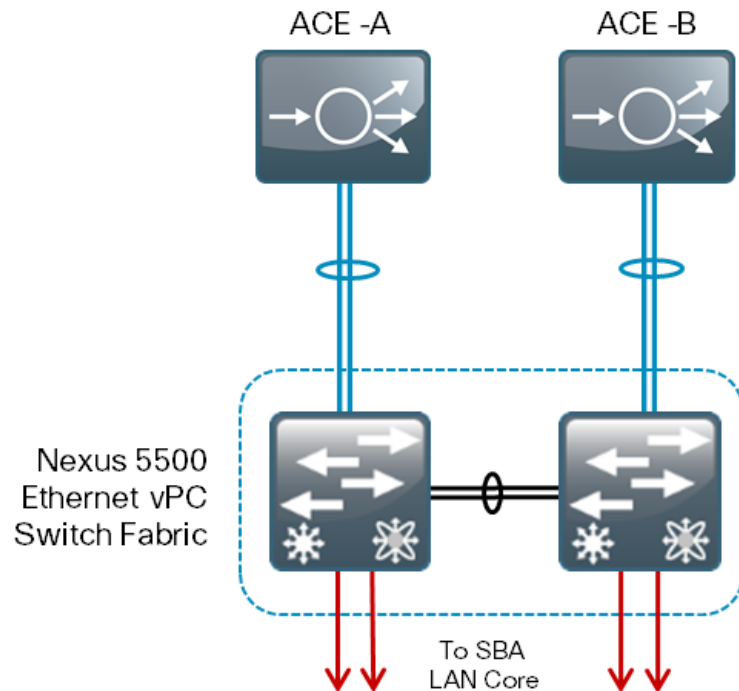
Figure 1 - Cisco ACE appliances deployed in a pair



Connectivity Option 1: Nexus 5500 and Virtual Port Channel (vPC) pair

The Cisco SBA data-center core uses the Cisco Nexus 5500UP switches and Virtual Port Channel (vPC) to connect many EtherChannel devices. For this type of deployment, follow Procedure 1, in the Configuring Cisco ACE Connectivity to Switches section.

Figure 2 - Nexus 5500 and vPC pair

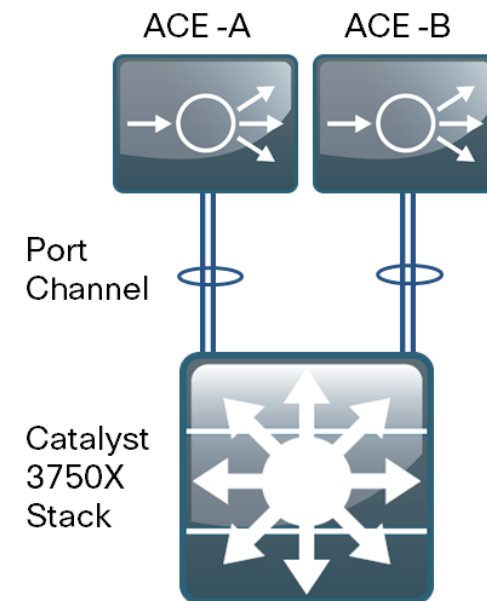


Connectivity Option 2: Cisco Catalyst 3750X Switch Stack

There are a number of places outside of the data center where a load balancing solution can be beneficial; for example, any place where there are servers that need a solution for high availability or scalability. This guide explains how to configure a Cisco 3750X stack, whether it resides in a server room or Internet DMZ.

As illustrated in Figure 3, Cisco ACE A should be connected to one switch in the stack and Cisco ACE B connected to the other. Even though the 3750X stack operates differently than the vPC stack, and has a data backplane between the switches in the stack, if the Cisco ACE appliance is cabled across multiple switches and a switch is taken out of service, then the available bandwidth to that appliance is reduced. It is better to fail over to the standby Cisco ACE appliance on a still functioning switch, than to run in a degraded mode.

Figure 3 - Cisco Catalyst 3750x switch stack



Process

Configuring Cisco ACE Connectivity to Switches

1. Configure port channels

Procedure 1

Configure port channels

Option 1. Connect port channels to Cisco Nexus 5500

With this option, you configure physical interfaces to the port channel on the Cisco Nexus 5500UP data center core-switch. If the vPC peer-link between the data center core-switches fails, then one of the switches will go into error recovery mode. This shuts down interfaces associated with VLANs that are part of vPC connections in order to prevent any loops in the infrastructure. Because the Cisco ACE appliances are single-homed to each data center core-switch, and use a VLAN that is part of other vPC connections, rather than a vPC for connectivity, they are non-vPC ports, or vPC orphan-ports.

During this configuration, you enter the **vpc orphan-port suspend** command. This command shuts down the EtherChannel interfaces to the attached Cisco ACE appliance on each switch in the event that the vPC peer link is broken between the data center core switches and a switch goes into error recovery mode. The active appliance on the switch that remains in service will continue operating and provides the resiliency in the design.

Cisco ACE supports EtherChannel, but does not support Link Aggregation Control Protocol (LACP); therefore, the **channel-group mode** will be forced on.

Step 1: Set the ports connected to the Cisco ACE appliance by using the **speed 1000** command. You can use the default of 10-Gigabit Ethernet, or choose a size down to 1-Gigabit Ethernet.



Tech Tip

When configuring the interfaces, the **vpc orphan-port suspend** command must be entered before the **channel-group** command. If you enter the **channel-group** command on the interface first, the switch will not let you enter the **vpc orphan-port suspend** command on the interface.

```
interface Ethernet1/3
  description ACE 1 Gig 1/1
  speed 1000
  vpc orphan-port suspend
  channel-group 13 mode on
```

```
interface Ethernet1/4
  description ACE 1 Gig 1/2
  speed 1000
  vpc orphan-port suspend
  channel-group 13 mode on
```

When you assign the channel-group to a physical interface, it creates the logical EtherChannel (port-channel) interface. In the next step, configure the logical port-channel interfaces on both data center core switches. The physical interfaces tied to the port-channel will inherit the settings.

Step 2: Configure the logical port-channel interfaces.

```
interface port-channel13
  switchport mode trunk
  switchport trunk allowed vlan 149,912
  spanning-tree port type edge trunk
```

Step 3: Configure an unused VLAN for the Cisco ACE fault tolerant heartbeat VLAN.

```
vlan 912
  name ACE-Heartbeat
```

Step 4: Apply the following configuration to the second Cisco Nexus 5500UP data center core-switch.

```
interface Ethernet1/3
  description ACE 2 Gig 1/1
  speed 1000
  vpc orphan-port suspend
  channel-group 13 mode on

interface Ethernet1/4
  description ACE 2 Gig 1/2
  speed 1000
  vpc orphan-port suspend
  channel-group 13 mode on

interface port-channel13
  switchport mode trunk
  switchport trunk allowed vlan 149,912
  spanning-tree port type edge trunk
  vlan 912
  name ACE-Heartbeat
```

Option 2. Connect port channels to Cisco Catalyst 3750X

With this option, you configure physical interfaces to the port channel on the Cisco Catalyst 3750X.

Step 1: Configure physical interfaces to the port channel on the Cisco Catalyst 3750X switch stack for Cisco ACE A. In a switch stack, you can use one port channel per Cisco ACE appliance.

```
interface GigabitEthernet1/0/45
  description ace4710-A g1/1
  channel-group 45 mode on
!
interface GigabitEthernet1/0/46
  description ace4710-A g1/2
  channel-group 45 mode on
```

Step 2: Configure physical interfaces to the port channel on the Cisco Catalyst 3750X switch stack for Cisco ACE B.

```
interface GigabitEthernet2/0/45
  description ace4710-B g1/1
  channel-group 46 mode on
!
interface GigabitEthernet2/0/46
  description ace4710-B g1/2
  channel-group 46 mode on
```

Step 3: Configure the logical port-channel interface for Cisco ACE A.

```
interface Port-channel45
  description ACE-A
  switchport trunk encapsulation dot1q
  switchport trunk allowed vlan 148,912
  switchport mode trunk
```

Step 4: Configure the logical port-channel interface for Cisco ACE B.

```
interface Port-channel46
  description ACE
  switchport trunk encapsulation dot1q
  switchport trunk allowed vlan 148,912
  switchport mode trunk
```

Process

Configuring Cisco ACE Appliance Network

1. Perform initial setup
2. Configure high availability

Procedure 1 Perform initial setup

Step 1: Connect to the Cisco ACE appliance via the console, perform the initial configuration, and then exit from the initial configuration dialog box at the prompt.

```
switch login: admin
Password: admin
Admin user is allowed to log in only from console until the
default password is changed.
www user is allowed to log in only after the default password
is changed.
Enter the new password for user "admin": password
Confirm the new password for user "admin": password
admin user password successfully changed.
Enter the new password for user "www": password
Confirm the new password for user "www": password
www user password successfully changed.
<text wall removed>
ACE>Would you like to enter the basic configuration dialog
(yes/no) [y]: n
switch/Admin#
```

Step 2: In config mode, set the system hostname.

```
hostname ACE4710-A
```

Step 3: Configure basic network security policies. This allows for management access to the Cisco ACE appliance.

```
access-list ALL line 8 extended permit ip any any
class-map type management match-any remote_access
  2 match protocol xml-https any
  3 match protocol icmp any
  4 match protocol telnet any
  5 match protocol ssh any
  6 match protocol http any
  7 match protocol https any
  8 match protocol snmp any
policy-map type management first-match remote_mgmt_allow_
policy
  class remote_access
    permit
```

Step 4: Configure port channel and trunking on the gigabit Ethernet interfaces.

```
interface gigabitEthernet 1/1
  channel-group 1
  no shutdown
interface gigabitEthernet 1/2
  channel-group 1
  no shutdown
interface port-channel 1
  switchport trunk native vlan 1
  switchport trunk allowed vlan 149
  no shutdown
```

With this configuration, a 2-Gbps port channel is provisioned, which is sufficient for a Cisco ACE 4710 appliance with up to a 2-Gbps license. If a 4-Gbps license is being used, include gigabit Ethernet ports 1/3 and 1/4 for a total of 4 Gbps of throughput.

Step 5: Configure the VLAN 149 interface on Cisco ACE for management access and general network connectivity.

```
interface vlan 149
  ip address 10.4.49.119 255.255.255.0
  access-group input ALL
  service-policy input remote_mgmt_allow_policy
  no shutdown
```

Step 6: Configure the default route.

```
ip route 0.0.0.0 0.0.0.0 10.4.49.1
```

Step 7: Configure NTP.

```
ntp server 10.4.48.17
```

Step 8: Configure SNMP.

```
snmp-server community cisco ro
```

You can now reach Cisco ACE 4710-A via the network.

To configure Cisco ACE 4710-B, repeat Step 1 through Step 6, replacing the hostname in Step 2 with ACE4710-B and the IP address in Step 5 with 10.4.49.120.

Procedure 2 Configure high availability

Next, you configure the Cisco ACE appliances as an active standby failover pair. After you configure high availability, the devices are synchronized and further configuration is necessary only on the primary appliance. Start with the Cisco ACE appliance that you want to be primary. In this example, the primary appliance is 10.4.49.119.

Step 1: Open a browser window and enter <https://10.4.49.119> into the address field. The Cisco ACE GUI opens.

Step 2: In the **Username** box, type **admin**, in the **Password** box, type the password that you configured in Procedure 1, Step 1, and then click **Log In**.



Step 3: Navigate to **Config > Virtual Contexts > High Availability (HA) > Setup**, and then click **Edit**.

Step 4: On the ACE HA Management dialog box, enter the following values, and then click **Deploy Now**.

- VLAN—**912**
- Interface—**Port Channel 1**
- IP Address—**10.255.255.1**
- IP Address Peer Appliance—**10.255.255.2**
- Netmask—**255.255.255.0**
- Management IP Address—**10.4.49.119**
- Management IP Address Peer Appliance—**10.4.49.120**

Step 5: On the ACE HA Groups dialog box, click **Add**.

Step 6: Leave all of the values at their defaults, and then click **Deploy Now**.

ACE HA Groups										
Group#	Enabled	Context	Priority (Actual)	Preempt	State	Peer Priority (Actual)	Peer State	Autosync Run	Autosync Startup	# Trackers
	<input checked="" type="checkbox"/>	Admin	100	<input checked="" type="checkbox"/>		100		<input checked="" type="checkbox"/>	<input type="checkbox"/>	
<div>Deploy Now Cancel</div>										

High availability is now configured on the primary Cisco ACE appliance. To configure high availability on the secondary appliance, you must log in to the secondary Cisco ACE appliance.

Step 7: Open a browser window and enter <https://10.4.49.120> into the address field. The Cisco ACE GUI opens.

Step 8: In the **Username** box, type **admin**, in the **Password** box, type the password that you configured in Procedure 1, Step 1, and then click **Log In**.

Step 9: Navigate to **Config > Virtual Contexts > High Availability (HA) > Setup**, and then click **Edit**.

Config > Virtual Contexts > High Availability (HA) > Setup

ACE HA Management

VLAN *	This Appliance	Peer Appliance
Interface *	Port Channel: 1	
IP Address *	10.255.255.2	10.255.255.1
Netmask *	255.255.255.0	
Management IP Address *	10.4.49.120	10.4.49.119
Query VLAN:	149	II/A
Heartbeat Count *	10	
Heartbeat Interval *	300	
Interface Enabled:	<input checked="" type="checkbox"/>	
Shared VLAN Host Id:		
Peer Shared VLAN Host Id:		
HA State:	TL Setup	

Deploy Now Cancel

ACE HA Groups

Group#	Enabled	Context	Priority (Actual)	Preempt	State	Peer Priority (Actual)	Peer State	Autosync Run	Autosync Startup	# Trackers
No records										
<div>Add Edit Switchover Delete Details...</div>										

Step 10: On the ACE HA Management dialog box, enter the following values, and then click **Deploy Now**.

- VLAN—**912**
- Interface—**Port Channel 1**
- IP Address—**10.255.255.2**
- IP Address Peer Appliance—**10.255.255.1**
- Netmask—**255.255.255.0**
- Management IP Address—**10.4.49.120**
- Management IP Address Peer Appliance—**10.4.49.119**

Step 11: On the ACE HA Groups dialog box, click **Add**.

Step 12: Leave all of the values at their defaults, and then click **Deploy Now**.

ACE HA Groups										
Group#	Enabled	Context	Priority (Actual)	Preempt	State	Peer Priority (Actual)	Peer State	Autosync Run	Autosync Startup	# Trackers
	<input checked="" type="checkbox"/>	Admin	100	<input checked="" type="checkbox"/>		100		<input checked="" type="checkbox"/>	<input type="checkbox"/>	
<div>Deploy Now Cancel</div>										

The two Cisco ACE appliances should be communicating and high availability should be up and active. The device you just finished configuring should show a state of Standby Hot and the peer should be Active, as illustrated below.

ACE HA Groups										
Group#	Enabled	Context	Priority (Actual)	Preempt	State	Peer Priority (Actual)	Peer State	Autosync Run	Autosync Startup	# Trackers
1	<input checked="" type="checkbox"/>	Admin	100 (100)	<input checked="" type="checkbox"/>	Standby Hot	100 (100)	Active	<input checked="" type="checkbox"/>	<input type="checkbox"/>	0
<div>Add Edit Switchover Delete Details...</div>										

Make any additional configurations on the primary Cisco ACE appliance, 10.4.49.119. All changes are automatically replicated to the secondary Cisco ACE appliance, 10.4.49.120.

Process

Setting up Load Balancing for HTTP Servers

1. Configure health probes
2. Configure real servers
3. Configure a server farm
4. Configure inband-health checking
5. Configure Remove mode
6. Configure a NAT pool
7. Configure a virtual server

Procedure 1 Configure health probes

Health probes poll the servers or applications to make sure that the server or service is available, and to allow the system to remove failed devices. In this procedure, you will build an Internet Control Message Protocol (ICMP) and an HTTP probe.

Step 1: Open a browser window and enter <https://10.4.49.119> into the address field. The Cisco ACE GUI opens.

Step 2: In the **Username** box, type [admin](#), in the **Password** box, type the password you configured in Procedure 1, Step 1, and then click **Log In**.

Step 3: Navigate to **Config > Virtual Contexts > Load Balancing > Health Monitoring**, and then click **Add**.

Step 4: On the New Health Monitoring dialog box, in the **Name** box, enter [icmp-probe](#), and in the **Type** list, choose **ICMP**.

Step 5: Click **Deploy Now**.

The screenshot shows the 'New Health Monitoring' dialog box. The 'Name' field contains 'icmp-probe'. The 'Type' dropdown is set to 'ICMP'. The 'Description' field is empty. The 'Probe Interval (Seconds)' is 15, 'Pass Detect Interval (Seconds)' is 60, and 'Fail Detect' is 3. There is a 'More Settings' link. At the bottom right, there are 'Deploy Now', 'Cancel', and a navigation arrow button.

Step 6: Navigate to **Config > Virtual Contexts > Load Balancing > Health Monitoring**, and then click **Add**.

Step 7: On the New Health Monitoring dialog box, in the **Name** box, enter [http-probe](#), and in the **Type** list, choose **HTTP**.

Step 8: Click **Deploy Now**.

The screenshot shows the 'New Health Monitoring' dialog box. The 'Name' field contains 'http-probe'. The 'Type' dropdown is set to 'HTTP'. The 'Description' field is empty. The 'Probe Interval (Seconds)' is 15, 'Pass Detect Interval (Seconds)' is 60, and 'Fail Detect' is 3. The 'Port' field is empty. The 'Request Method Type' has radio buttons for 'Get' and 'Head', with 'Head' selected. The 'Request HTTP URL' field is empty. There is a 'More Settings' link. At the bottom right, there are 'Deploy Now', 'Cancel', and a navigation arrow button.

Step 9: On the Expect Status tab, click **Add**.

Step 10: For both the maximum and minimum status codes, enter [200](#), and then click **Deploy Now**.

The screenshot shows the 'Expect Status' tab. The title is 'New Expect Status @ http-probe'. There are two input fields: 'Min. Expect Status Code' and 'Max. Expect Status Code', both containing the value '200'. At the bottom right, there are 'Deploy Now', 'Cancel', and a navigation arrow button.

You have just created the ICMP and HTTP probes, which you can use to monitor the real and virtual servers in the load balancing server farm.

Procedure 2 Configure real servers

In this procedure you will add the real servers, across which the Cisco ACE appliance will load balance client connections.

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Real Servers**, and then click **Add**.

Step 2: On the New Real Server dialog box, enter the following values, and then click **Deploy Now**.

- Name—**webserver1**
- IP Address—**10.4.49.111**
- Probes—**icmp-probe**

The screenshot shows the 'New Real Server' dialog box. The 'Name' field is set to 'webserver1'. The 'Type' is 'Host' and 'Redirect' is unselected. The 'State' is 'In Service' and 'Out of Service' is unselected. The 'Description' field is empty. The 'IP Address Type' is 'IPv4' and 'IPv6' is unselected. The 'IPv4 Address' is '10.4.49.111'. The 'Fail-On-All' checkbox is unchecked. The 'Min. Connections' is '4000000' and 'Max. Connections' is '4000000'. The 'Weight' is '8'. The 'Probes' section shows 'Available' with 'http-probe' and 'Selected' with 'icmp-probe'. The 'Rate Bandwidth' and 'Rate Connection' fields are empty. At the bottom are 'Deploy Now', 'Cancel', and '>' buttons.

Step 3: Navigate to **Config > Virtual Contexts > Load Balancing > Real Servers**, and then click **Add**.

Step 4: On the New Real Server dialog box, enter the following values, and then click **Deploy Now**.

- Name—**webserver2**
- IP Address—**10.4.49.112**
- Probes—**icmp-probe**

This example uses the ICMP probe to monitor the real servers configured in this example, thereby ensuring the server is monitored rather than a specific service. This is the most flexible configuration and allows load-balancing for multiple services on a single physical or virtual server.

You have just configured the two web servers. If you have additional servers that you plan on using, you can configure them now by repeating Procedure 2.

Procedure 3 Configure a server farm

A server farm on the Cisco ACE appliance is a pool of real servers that you can use to connect to the virtual IP address that the clients will use to connect to the HTTP service.

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Server Farms**, and then click **Add**.

Step 2: On the New Server Farm dialog box, enter the following values, and then click **Deploy Now**.

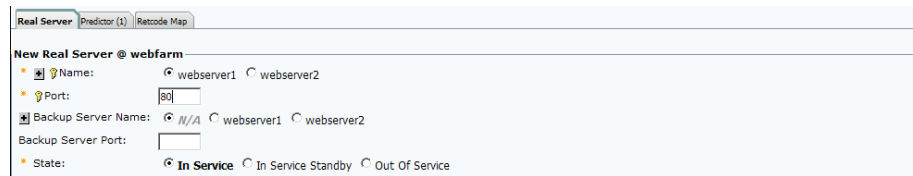
- Name—**webfarm**
- Probes—**http-probe**

The screenshot shows the 'New Server Farm' dialog box. The 'Name' field is set to 'webfarm'. The 'Type' is 'Host' and 'Redirect' is unselected. The 'Description' field is empty. The 'Fail Action' is 'N/A', 'Purge', and 'Reassign' are unselected. The 'Fail-On-All' checkbox is unchecked. The 'Inband-Health Check' dropdown is set to 'N/A'. The 'Transparent' checkbox is unchecked. The 'Dynamic Workload Scaling' is 'N/A', 'Burst', and 'Local' are unselected. The 'Partial-Threshold Percentage' is '0'. The 'Back Inservice' is '0'. The 'Probes' section shows 'Available' with 'icmp-probe' and 'Selected' with 'http-probe'. At the bottom are 'Deploy Now', 'Cancel', and '>' buttons.

Step 3: On the Real Server tab, click **Add**.

Step 4: On the New Real Server dialog box, next to **Name**, select **web-server1**, and then in the **Port** box, enter **80** for HTTP.

Step 5: Click **Deploy Now**.



Step 6: On the Real Server tab, click **Add**.

Step 7: On the New Real Server dialog box, next to **Name**, select **webserver2**, and then in the **Port** box, enter **80**.

Step 8: Click **Deploy Now**.

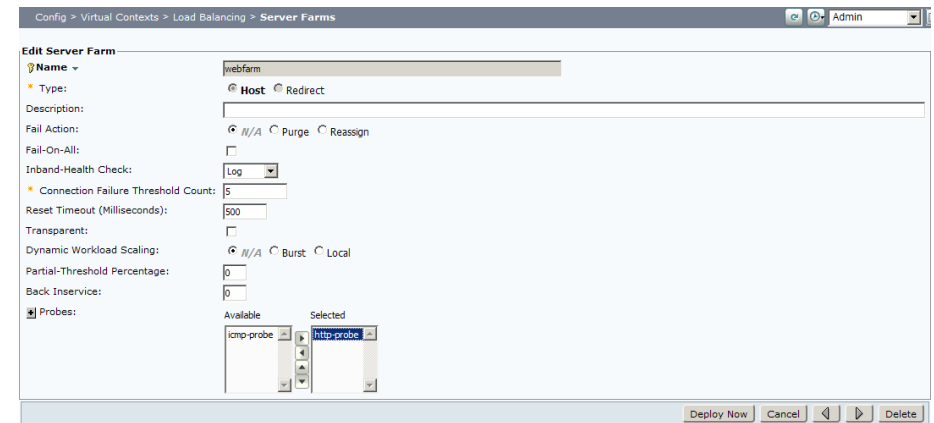
Step 9: On the Edit Server Farm dialog box, click **Deploy Now**.

You have just created the server-farm, webfarm, with the real-server members, webserver1 and webserver2, for HTTP on port 80. The http-probe will monitor all of the servers in the server farm to ensure that the HTTP service is available.

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Server Farms**, select **webfarm**, and then click **View/Edit**.

Step 2: On the Edit Server Farm dialog box, enter the following values, and then click **Deploy Now**.

- Inband-Health Check—**Log**
- Connection Failure Threshold Count—**5**
- Reset Timeout (Milliseconds)—**500**



Servers in the webfarm are now being monitored for TCP errors. If five errors occur within a 500-ms period, a syslog message will be sent to the NMS. If there is not a syslog server available on the network, the inband-health check can be set to use Count mode, and local statistics will be maintained on Cisco ACE and can be checked from the CLI.

Step 3: At the bottom of the **Server Farms** window, click the **Retcode Map** tab, and then click **Add**.

Procedure 4 Configure inband-health checking

Inband-health checking on Cisco ACE monitors return traffic and looks for failures from the real servers to the clients. It can identify, faster than active probes, when a server is having issues. When a failure is detected, the following modes are available:

- **Count**—Logs the failures locally on Cisco ACE, allowing you to view server issues from the CLI.
- **Log**—Triggers a syslog message to be sent to a Network Management System (NMS), as well as keeping the log locally on Cisco ACE.
- **Remove**—Triggers a log and takes the server out of service.

In this procedure, Log mode is used. This is because a small amount of errors of this type are normal on servers. Without more information about the server farm, using Remove mode could mean that the threshold would be too low and would take a system out of service unnecessarily, or too high and not take a failing server out of service. Log mode allows you to see errors and identify which real sever is having problems. Procedure 5, below, is an optional procedure that gives you guidance on using Remove mode.

Step 4: On the New Retcode Map dialog box, enter the following values, and then click **Deploy Now**.

- Lowest Retcode—**404**
- Highest Retcode—**404**
- **Type**—**Log**
- **Threshold**—**5**
- **Reset**—**10**

The screenshot shows the 'New Retcode Map @ webfarm' dialog box. It has tabs for 'Real Server', 'Predictor', and 'Retcode Map'. The 'Retcode Map' tab is active. The fields are: Lowest Retcode: 404, Highest Retcode: 404, Type: Log (selected), Threshold: 5, and Reset (Seconds): 10. At the bottom, there are buttons for 'Deploy Now', 'Cancel', and a right arrow.

If, within a 10-second period, a server in the webfarm responds to a client five times with the HTTP return code 404, a syslog message will be sent to the NMS.

Step 5: At the bottom of the **Server Farms** window, click the **Retcode Map** tab, and then click **Add**.

Step 6: On the New Retcode Map dialog box, enter the following values, and then click **Deploy Now**.

- Lowest Retcode—**500**
- Highest Retcode—**505**
- **Type**—**Log**
- **Threshold**—**5**
- **Reset**—**10**

The screenshot shows the 'New Retcode Map @ webfarm' dialog box. It has tabs for 'Real Server', 'Predictor', and 'Retcode Map'. The 'Retcode Map' tab is active. The fields are: Lowest Retcode: 500, Highest Retcode: 505, Type: Log (selected), Threshold: 5, and Reset (Seconds): 10. At the bottom, there are buttons for 'Deploy Now', 'Cancel', and a right arrow.

If, within a 10-second period, a server in the webfarm responds to a client five times with the HTTP return code in the range of 500 to 505, a syslog message will be sent to the NMS.

Step 7: Navigate to **Config > Virtual Contexts > System > Syslog**, and then check **Enable Syslog**.

The screenshot shows the 'Syslog' configuration page under 'Config > Virtual Contexts > System'. The 'Enable Syslog' checkbox is checked. Other settings include Facility: 20, Buffered Level, Console Level, History Level, Monitor Level, Persistence Level, Trap Level: 5-Notification, Queue Size, Enable Timestamp, Enable Standby, Enable Fastpath Logging, and Device Id Type: Undefined. At the bottom right, there is a 'Deploy Now' button.

Step 8: On the **Log Host** tab, click **Add**, enter **10.4.48.35**, and then click **Deploy Now**.

The screenshot shows the 'New Log Host' dialog box. It has tabs for 'Log Host', 'Log Message', and 'Log Rate Limit'. The 'Log Host' tab is active. The fields are: IP Address: 10.4.48.35, Protocol: UDP (selected), Protocol Port: 514, and Format: Emblem. At the bottom right, there are buttons for 'Deploy Now', 'Cancel', and a right arrow.

Step 9: On the Syslog dialog box, click **Deploy Now**.

Now the syslog messages that are triggered by the inband-health checks are sent to the syslog server at 10.4.48.35.

Procedure 5 **Configure Remove mode**

(Optional)

By following this procedure, you can configure inband-health checking to remove malfunctioning servers from the server farm. Before the inband-health checking is set to remove servers, there is some information that must be gathered about the server farm. The feature counts the number of errors over a specific interval and will remove a server that exceeds the configured threshold. As discussed in Procedure 4, a certain number of TCP and HTTP error is normal, so care must be taken not to set the threshold too low and cause a healthy server to be taken out of service.

If not already configured, use Procedure 4, above, to set up inband-health checking in either Log or Count mode for the server farm, webfarm.

Step 1: Connect to the primary ACE appliance CLI and enter the `sh server-farm webfarm` command. You should see a connection count and failure number per real server.

```
serverfarm      : webfarm, type: HOST
total rservers  : 2
state           : ACTIVE
DWS state       : DISABLED
```

```
-----connections-----
real      weight state  current  total  failures
---+-----+-----+-----+-----+-----+
rserver: webserver1
10.4.49.111:80    8  OPERATIONAL    0      2      0
rserver: webserver2
10.4.49.112:80    8  OPERATIONAL    0      2      0
```

Monitor this information for the duration of a normal business cycle for the application on the servers; this could range from a day to several weeks. From this information, you should be able to assess the percentage of failed connections for the set of real-servers in the server farm. This is your base-line error rate. You also need to know the average connections per second (CPS) for the servers in the server farm.

To configure Remove mode in inband-health checking, you need to determine the error threshold, monitoring interval, and resume-service interval. In this guide, a 500 CPS and a normal error-rate of 0.05% is used, and a server is taken out of service when the connection error rate reaches 0.5%. The

default sample rate for health checking on Cisco ACE is 100 ms. You need to calculate the number of errors needed to reach 0.5%. If the connection rate is 500 CPS, then at a 100-ms sample rate, you will see 50 connections per interval. A 0.5% error-rate for 50 connections is only .25 connections. The minimum threshold you can set is one, which gives you a 2% error-rate, which still does not hit your target of 0.5%. The easy way to hit the target error-rate is to increase the sample rate to 400 ms. A threshold of one would hit an error rate of 0.5%, but if you do this a single error will take the server out of service. Some errors are normal so you don't want to do this. Instead, you can increase the interval to 1000-ms and set the threshold to three. With this configuration, you are at a 0.6% error rate, which is very close to your target.

Step 2: Navigate to **Config > Virtual Contexts > Load Balancing > Server Farms**, select **webfarm** click **View/Edit**.

Step 3: On the Edit Server Farm dialog box enter the following values and then click **Deploy Now**:

- Inband-Health Check—**Remove**
- Connection Failure Threshold Count—**3**
- Reset Timeout (Milliseconds)—**1000**
- Resume Service (Seconds)—**300**

The screenshot shows the 'Edit Server Farm' configuration page for 'webfarm'. The page is titled 'Config > Virtual Contexts > Load Balancing > Server Farms'. The 'Name' field is 'webfarm'. The 'Type' is 'Host'. The 'Description' field is empty. The 'Fail Action' is 'N/A'. The 'Fail-On-All' checkbox is unchecked. The 'Inband-Health Check' is set to 'Remove'. The 'Connection Failure Threshold Count' is set to '3'. The 'Reset Timeout (Milliseconds)' is set to '1000'. The 'Resume Service (Seconds)' is set to '300'.

With this configuration, if three errors are received in one-second, then the server will be taken out of service. Resume is set to 300-seconds, so the server will go back into service after five minutes. If you want the server to stay out of service until someone manually intervenes, then do not set a value for resume.



Tech Tip

Cisco ACE uses reset = 100 ms as the default value. It is recommended that you use the default, unless the traffic to the VIP is very low.

Procedure 6

Configure a NAT pool

Step 1: Navigate to **Config > Virtual Contexts > Network > NAT Pools**, and then click **Add**.

Step 2: On the New NAT Pool dialog box, enter the following values, and then click **Deploy Now**.

- Start IP Address—**10.4.49.99**
- End IP Address—**10.4.49.99**
- Netmask—**255.255.255.0**

Procedure 7

Configure a virtual server

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Virtual Servers**, and then click **Add**.

Step 2: On the Properties dialog box, enter the following values:

- Virtual Server Name—**http-vip**
- Virtual IP Address—**10.4.49.100**
- VLAN—**149**

Step 3: On the Default L7 Load-Balancing Action dialog box, in the **Server Farm** list, choose **webfarm**, and then select **Deflate**.

Step 4: On the NAT dialog box, click **Add**, click **OK**, and then click **Deploy Now**.

Clients going to the virtual IP, 10.4.49.100 on port 80, will be load-balanced across the real servers, webserver1 and webserver2, in the server farm, webfarm.

Process

Load-Balancing and SSL-Offloading for HTTPS Servers

1. Configure real servers
2. Configure a server farm
3. Configure SSL proxy service
4. Configure HTTP-cookie sticky service
5. Configure a virtual server
6. Configure an HTTP to HTTPS Redirect

You can configure a group of servers for load balancing, in which the Cisco ACE appliance performs all of the SSL processing, thereby offloading it from the servers.

Procedure 1 Configure real servers

In this procedure you will add the real servers, across which the Cisco ACE appliance will load-balance client SSL connections.

Step 1: Open a browser window and enter <https://10.4.49.119> into the address field. The Cisco ACE GUI opens.

Step 2: In the **Username** box, type **admin**. In the **Password** box, type the password you configured in Procedure 1, Step 1, and then click **Log In**.

Step 3: Navigate to **Config > Virtual Contexts > Load Balancing > Real Servers**, and then click **Add**.

Step 4: On the New Real Server dialog box, enter the following values, and then click **Deploy Now**.

- Name—**webserver3**
- IP Address—**10.4.49.113**
- Probes—**icmp-probe**

Step 5: Navigate to **Config > Virtual Contexts > Load Balancing > Real Servers**, and then click **Add**.

Step 6: On the New Real Server dialog box, enter the following values, and then click **Deploy Now**.

- Name—**webserver4**
- IP Address—**10.4.49.114**
- Probes—**icmp-probe**

In this example, the ICMP-probe monitors the real servers, thereby ensuring that the server is monitored, rather than a specific service. This is the most flexible configuration and allows load-balancing for multiple services on a single physical or virtual server.

You have just configured the two web servers. If you have additional servers that you plan on using, you can configure them now by repeating Procedure 1.

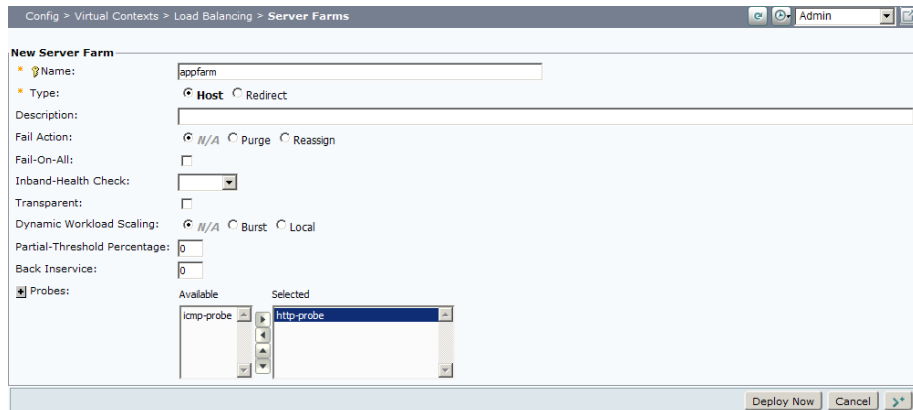
Procedure 2 Configure a server farm

A server farm is a pool of real servers that you can use to connect to the VIP-address that the clients will use to connect to the HTTP service.

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Server Farms**, and then click **Add**.

Step 2: On the New Server Farm dialog box, enter the following values, and then click **Deploy Now**.

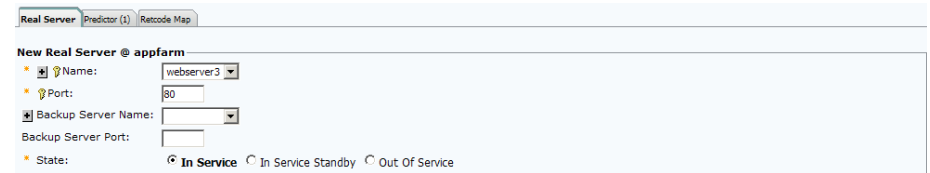
- Name—**appfarm**
- Probes—**http-probe**



Step 3: On the Real Server tab, click **Add**.

Step 4: On the New Real Server dialog box, in the **Name** list, choose **webserver3**, and then in the **Port** box, enter **80** for HTTP.

Step 5: Click **Deploy Now**.



Step 6: Click **Deploy Now**.

Step 7: On the Real Server tab, click **Add**.

Step 8: On the New Real Server dialog box, in the **Name** list, choose **webserver4**, and then in the **Port** box, enter **80**.

Step 9: Click **Deploy Now**.

Step 10: On the Edit Server Farm dialog box, click **Deploy Now**.

You have just created the server farm, appfarm, with the real-server members, webserver3 and webserver4, for HTTP on port 80. The Cisco ACE appliance will perform all of the SSL-processing so, even though clients will access the application on these servers via HTTPS, the traffic from Cisco ACE to the servers will happen over port 80. The http-probe will monitor all of the servers in the server farm to ensure that the HTTP service is available.

Procedure 3 Configure SSL proxy service

In order for Cisco ACE to offload the SSL processing, you need to configure an SSL proxy service. In this guide, the Cisco sample key and certificate is used. However, in a production deployment, you would most likely purchase a certificate from a trusted certificate authority (CA).

Step 1: Navigate to **Config > Virtual Contexts > SSL > Proxy Service**, and then click **Add**.

Step 2: On the New Proxy Service dialog box, in the **Name** box, enter **app-ssl-proxy**.

Step 3: Select both **cisco-sample-key** and **cisco-sample-cert**, and then click **Deploy Now**.

Step 5: Next to Sticky Server Farm, select **appfarm**, and then click **Deploy Now**.

Procedure 4 Configure HTTP-cookie sticky service

The HTTP-cookie sticky service keeps traffic from a client stuck to a single real-server. This is useful when state could be lost if the client connection was balanced across several servers.

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Stickiness**, and then click **Add**.

Step 2: On the New Sticky Group dialog box, in the **Group Name** box, enter **app-sticky**.

Step 3: In the **Type** list, choose **HTTP Cookie**, and then in the **Cookie Name** box, enter **APPSSESSIONID**.

Step 4: Select both **Enable Insert** and **Browser Expire**.

Procedure 5 Configure a virtual server

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Virtual Servers**, and then click **Add**.

Step 2: On the Properties dialog box, enter the following values:

- Virtual Server Name—**https-vip**
- Virtual IP Address—**10.4.49.101**
- Application Protocol—**HTTPS**
- VLAN—**149**

Step 3: On the SSL Termination dialog box, in the **Proxy Service Name** list, choose **app-ssl-proxy**.

Step 4: On the Default L7 Load-Balancing Action dialog box, in the **Primary Action** list, choose **Sticky**.

Step 5: In the **Sticky Group** list, choose **app-sticky (HTTP Cookie)**, and then select **Deflate**.

The image shows two stacked configuration dialog boxes. The top box is titled "SSL Termination" and has a "Proxy Service Name" dropdown set to "app-ssl-proxy" with a "View" button next to it. The bottom box is titled "Default L7 Load-Balancing Action". It has an "Action:" section with "Primary Action:" set to "Sticky" and "Sticky Group:" set to "app-sticky (HTTP Cookie)" with a "View" button. Below this, "Compression Method:" has radio buttons for "Deflate" (selected), "Gzip", and "N/A". A note states: "Exclude the following MIME Types from HTTP compression: .gif, .css, .js, .class, .jar, .cab, .txt, .ps, .vbs, .xsl, .xml, .pdf, .swf, .jpg, .jpeg, .jpe, .png". At the bottom, "SSL Initiation:" has a dropdown menu.

Step 6: On the NAT dialog box, click **Add**, click **OK**, and then click **Deploy Now**.

The image shows the "NAT" configuration dialog box. It has a yellow warning banner at the top that says "Source NAT needs to be configured for IPv4-to-IPv6 / IPv6-to-IPv4 virtual server, otherwise it will not be functional." with a link "How to add a NAT Pool ID". Below the banner is a table with two columns: "VLAN" and "NAT Pool ID (Begin IP - End IP: Netmask: PAT)". The first row shows "149 (1 Pools Available)" and "1 (10.4.49.99 - 10.4.49.99: 255.255.255.0: PAT Enabled)". At the bottom right are "OK" and "Cancel" buttons. At the bottom center are "Deploy Now" and "Cancel" buttons.

Clients going to the virtual IP, 10.4.49.101 on port 443, will be load-balanced across the real-servers, webserver3 and webserver4, in the server farm, appfarm. Cisco ACE will terminate the SSL session and load-balance the connections to the real-servers over standard HTTP on TCP port 80.

Procedure 6

Configure an HTTP to HTTPS Redirect

(Optional)

It is often preferable to have HTTP traffic redirected to HTTPS to ensure that connections to that service are encrypted. By following this procedure, you can create a service that redirects any HTTP traffic directed to 10.4.49.101 to the HTTPS service configured above.

Step 1: Navigate to **Config > Virtual Contexts > Load Balancing > Real Servers**, and then click **Add**.

Step 2: On the New Real Server dialog box, enter the following values, and then click **Deploy Now**.

- Name—**redirect1**
- Type—**Redirect**
- Web Host Redirection—**https://%h%p**
- Redirection Code—**302**

The image shows the "New Real Server" configuration dialog box. The "Name" field is set to "redirect1". The "Type:" section has radio buttons for "Host" and "Redirect" (selected). The "State:" section has radio buttons for "In Service" (selected) and "Out Of Service". The "Description:" field is empty. The "Min. Connections:" and "Max. Connections:" fields are both set to "4000000". The "Web Host Redirection:" field is set to "https://%h%p". The "Redirection Code:" section has radio buttons for "N/A", "301", and "302" (selected). The "Probes:" section has "Available" and "Selected" tabs, with "Selected" being active. The "Rate Bandwidth:" and "Rate Connection:" fields are empty. At the bottom right are "Deploy Now", "Cancel", and ">" buttons.

Step 3: Navigate to **Config > Virtual Contexts > Load Balancing > Server Farms**, and then click **Add**.

Step 4: On the New Server Farm dialog box, enter the following values, and then click **Deploy Now**.

- Name—**http-redirect**
- Type—**Redirect**

Step 5: On the Real Server tab, click **Add**.

Step 6: On the New Real Server dialog box, select **redirect1**, and then click **Deploy Now**.

Step 7: On the Edit Server Farm dialog box, click **Deploy Now**.

Step 8: Navigate to **Config > Virtual Contexts > Load Balancing > Virtual Servers**, and then click **Add**.

Step 9: On the Properties dialog box, enter the following values:

- Virtual Server Name—**http-vip-redirect**
- Virtual IP Address—**10.4.49.101**
- VLAN—**149**

Step 10: On the Default L7 Load-Balancing Action dialog box, in the **Server Farm** list, choose **http-redirect**, and then click **Deploy Now**.

Appendix A: Product List

Data Center Services

Functional Area	Product Description	Part Numbers	Software
Application Resiliency	Cisco ACE 4710 Application Control Engine 2Gbps	ACE-4710-02-K9	A5(1.2)
	Cisco ACE 4710 Application Control Engine 1Gbps	ACE-4710-01-K9	
	Cisco ACE 4710 Application Control Engine 1Gbps 2-Pack	ACE-4710-2PAK	
	Cisco ACE 4710 Application Control Engine 500 Mbps	ACE-4710-0.5-K9	

Data Center Core

Functional Area	Product Description	Part Numbers	Software
Core Switch	Cisco Nexus 5596 up to 96-port 10GbE, FCoE, and Fibre Channel SFP+	N5K-C5596UP-FA	NX-OS 5.1(3)N1(1a) Layer 3 License
	Cisco Nexus 5596 Layer 3 Switching Module	N55-M160L30V2	
	Cisco Nexus 5548 up to 48-port 10GbE, FCoE, and Fibre Channel SFP+	N5K-C5548UP-FA	
	Cisco Nexus 5548 Layer 3 Switching Module	N55-D160L3	

Server Room

Functional Area	Product Description	Part Numbers	Software
Stackable Ethernet Switch	Cisco Catalyst 3750-X Series Stackable 48 Ethernet 10/100/1000 ports	WS-C3750X-48T-S	15.0(1)SE2
	Cisco Catalyst 3750-X Series Stackable 24 Ethernet 10/100/1000 ports	WS-C3750X-24T-S	IP Base
	Cisco Catalyst 3750-X Series Four GbE SFP ports network module	C3KX-NM-1G	

Appendix B: Configuration

This Cisco ACE 4710 appliance is one of a resilient pair that provide Layer 4 through Layer 7 switching services. The configuration below is the from the primary appliance of the pair of Cisco ACS 4710s. The configuration of the secondary Cisco ACE appliance is identical except that the hostname is unique per appliance and that the interface and failover IP address are reversed.

```
no ft auto-sync startup-config
```

```
logging enable
logging timestamp
logging trap 5
logging host 10.4.48.35 udp/514 format emblem
```

```
boot system image:c4710ace-t1k9-mz.A5_1_2.bin
```

```
peer hostname ACE4710-B
hostname ACE4710-A
interface gigabitEthernet 1/1
    channel-group 1
    no shutdown
interface gigabitEthernet 1/2
    channel-group 1
    no shutdown
interface gigabitEthernet 1/3
    shutdown
interface gigabitEthernet 1/4
    shutdown
interface port-channel 1
    ft-port vlan 912
    switchport trunk native vlan 1
```

```
switchport trunk allowed vlan 149
no shutdown
```

```
ntp server 10.4.48.17
```

```
access-list ALL line 8 extended permit ip any any
```

```
probe http http-probe
    request method head
    expect status 200 200
probe icmp icmp-probe
```

```
rserver redirect redirect1
    conn-limit max 4000000 min 4000000
    webhost-redirection https://%h%p 302
inservice
```

```
rserver host webserver1
    ip address 10.4.49.111
    conn-limit max 4000000 min 4000000
    probe icmp-probe
inservice
```

```
rserver host webserver2
    ip address 10.4.49.112
    conn-limit max 4000000 min 4000000
    probe icmp-probe
inservice
```

```
rserver host webserver3
    ip address 10.4.49.113
    conn-limit max 4000000 min 4000000
    probe icmp-probe
inservice
```

```
rserver host webserver4
    ip address 10.4.49.114
    conn-limit max 4000000 min 4000000
```

```

probe icmp-probe
inservice

serverfarm host appfarm
  probe http-probe
  rserver webserver3 80
    conn-limit max 4000000 min 4000000
    inservice
  rserver webserver4 80
    conn-limit max 4000000 min 4000000
    inservice
serverfarm redirect http-redirect
  rserver redirect1
    conn-limit max 4000000 min 4000000
    inservice
serverfarm host webfarm
  probe http-probe
  inband-health check remove 3 reset 1000 resume-service 300
  retcode 404 404 check log 5 reset 10
  retcode 500 505 check log 5 reset 10
  rserver webserver1 80
    conn-limit max 4000000 min 4000000
    inservice
  rserver webserver2 80
    conn-limit max 4000000 min 4000000
    inservice

sticky http-cookie APPSESSIONID app-sticky
  cookie insert browser-expire
  serverfarm appfarm

ssl-proxy service app-ssl-proxy
  key cisco-sample-key
  cert cisco-sample-cert

class-map type http loadbalance match-any default-compression-
exclusion-mime-type

```

```

description DM generated classmap for default LB compression
exclusion mime types.
  2 match http url .*gif
  3 match http url .*css
  4 match http url .*js
  5 match http url .*class
  6 match http url .*jar
  7 match http url .*cab
  8 match http url .*txt
  9 match http url .*ps
  10 match http url .*vbs
  11 match http url .*xsl
  12 match http url .*xml
  13 match http url .*pdf
  14 match http url .*swf
  15 match http url .*jpg
  16 match http url .*jpeg
  17 match http url .*jpe
  18 match http url .*png
class-map match-all http-vip
  2 match virtual-address 10.4.49.100 tcp eq www
class-map match-all http-vip-redirect
  2 match virtual-address 10.4.49.101 tcp eq www
class-map match-all https-vip
  2 match virtual-address 10.4.49.101 tcp eq https
class-map type management match-any remote_access
  2 match protocol xml-https any
  3 match protocol icmp any
  4 match protocol telnet any
  5 match protocol ssh any
  6 match protocol http any
  7 match protocol https any
  8 match protocol snmp any

policy-map type management first-match remote_mgmt_allow_policy
  class remote_access
    permit

```

```

policy-map type loadbalance first-match http-vip-l7slb
  class default-compression-exclusion-mime-type
    serverfarm webfarm
  class class-default
    serverfarm webfarm
    compress default-method deflate
policy-map type loadbalance first-match http-vip-redirect-l7slb
  class class-default
    serverfarm http-redirect
policy-map type loadbalance first-match https-vip-l7slb
  class default-compression-exclusion-mime-type
    sticky-serverfarm app-sticky
  class class-default
    compress default-method deflate
    sticky-serverfarm app-sticky

policy-map multi-match int149
  class http-vip
    loadbalance vip inservice
    loadbalance policy http-vip-l7slb
    nat dynamic 1 vlan 149
  class https-vip
    loadbalance vip inservice
    loadbalance policy https-vip-l7slb
    nat dynamic 1 vlan 149
    ssl-proxy server app-ssl-proxy
  class http-vip-redirect
    loadbalance vip inservice
    loadbalance policy http-vip-redirect-l7slb

interface vlan 149
  ip address 10.4.49.119 255.255.255.0
  peer ip address 10.4.49.120 255.255.255.0
  access-group input ALL
  nat-pool 1 10.4.49.99 10.4.49.99 netmask 255.255.255.0 pat
  service-policy input remote_mgmt_allow_policy
  service-policy input int149

```

```

no shutdown

ft interface vlan 912
  ip address 10.255.255.1 255.255.255.0
  peer ip address 10.255.255.2 255.255.255.0
  no shutdown

ft peer 1
  heartbeat interval 300
  heartbeat count 10
  ft-interface vlan 912
ft group 1
  peer 1
  associate-context Admin
  inservice

ip route 0.0.0.0 0.0.0.0 10.4.49.1

snmp-server community cisco group Network-Monitor

username admin password 5 ***** role Admin domain default-  

domain
username www password 5 ***** role Admin domain default-domain

```

Appendix C: Changes

This appendix summarizes the changes to this guide since the previous Cisco SBA series.

- We updated the Cisco ACE software to version A5(1.2)
- We added inband health checking for TCP and HTTP traffic
- We added details on connectivity to Cisco Nexus 5500 Series switches.

Notes

Feedback

Click [here](#) to provide feedback to Cisco SBA.



SMART BUSINESS ARCHITECTURE

ALL DESIGNS, SPECIFICATIONS, STATEMENTS, INFORMATION, AND RECOMMENDATIONS (COLLECTIVELY, "DESIGNS") IN THIS MANUAL ARE PRESENTED "AS IS," WITH ALL FAULTS. CISCO AND ITS SUPPLIERS DISCLAIM ALL WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE. IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THE DESIGNS, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THE DESIGNS ARE SUBJECT TO CHANGE WITHOUT NOTICE. USERS ARE SOLELY RESPONSIBLE FOR THEIR APPLICATION OF THE DESIGNS. THE DESIGNS DO NOT CONSTITUTE THE TECHNICAL OR OTHER PROFESSIONAL ADVICE OF CISCO, ITS SUPPLIERS OR PARTNERS. USERS SHOULD CONSULT THEIR OWN TECHNICAL ADVISORS BEFORE IMPLEMENTING THE DESIGNS. RESULTS MAY VARY DEPENDING ON FACTORS NOT TESTED BY CISCO.

Any Internet Protocol (IP) addresses used in this document are not intended to be actual addresses. Any examples, command display output, and figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses in illustrative content is unintentional and coincidental.

© 2012 Cisco Systems, Inc. All rights reserved.



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)