

Cisco Global Site Selector 4480



The Cisco GSS 4480 Global Site Selector delivers site selection services that are critical for any business resilience strategy to optimize multisite deployments involving globally distributed data centers.

The Cisco GSS 4480 is a networking product that globally load balances distributed data centers. The Cisco GSS 4480 acts as the cornerstone of multisite disaster recovery plans in deployments of Cisco's market-leading content switches. Customers deploying new Cisco content switches such as the Cisco CSS 11500 Content Services Switch and the Content Switching Module (CSM) for the Cisco Catalyst® 6500 Series switches or have already deployed legacy switches such as the Cisco CSS 11000 and Cisco Local Directors can benefit from the new levels of traffic management and centralized command and control provided by the Cisco GSS 4480.

The Cisco GSS 4480 product delivers the following key capabilities:

- Provides a scalable, dedicated hardware platform for Cisco's industry-leading content switches to ensure Web-based applications are always available, by detecting site outages or site congestion and rerouting content requests
- Improves global data center selection process by offering user-selectable global load-balance algorithms

- Offloads Domain Named System (DNS) servers by taking over the domain resolution process, and transmits these requests at thousands of requests per second
- Scales to support hundreds of data centers or server load balancers (SLBs)
- Complements existing DNS infrastructure by providing centralized domain management
- Tightly integrates with Cisco SLBs without sacrificing the ability to work in a heterogeneous environment of DNS-capable networking products

The Cisco GSS 4480 allows businesses to deploy global Internet and intranet applications with the confidence that Web application users will be quickly rerouted to a standby data center if a primary data-center outage or overload occurs. The Cisco GSS 4480 traffic management process continuously monitors the load and health of the SLBs within each data center. This information is used in conjunction with customer-controlled load-balancing algorithms to enable the Cisco GSS 4480 to—in real time—select a data center that is available and not overloaded within user-definable load conditions.



By offloading the DNS server resolution process from traditional DNS servers, the Cisco GSS 4480 can optimize global site selection, boost DNS responsiveness, assure data-center availability, and increase scalability of Web sites and data centers. The Cisco GSS 4480 is a critical component for enterprises and service providers deploying globally distributed data centers, or installing disaster recovery systems for Web-based applications.

Key Features and Benefits

- Provides resilient architecture critical for disaster recovery and multisite Web applications deployments
- Offers flexible heterogeneous support for all Cisco SLBs and DNS-capable networking products
- Provides centralized command and control of DNS resolution process for direct and precise control of global load-balancing process
- Augments and offloads any DNS Berkeley Internet Name Domain (BIND) infrastructure to optimize both content requests and delivery for all types of static and dynamic Web content
- Provides dedicated processing of DNS requests for greater performance and scalability
- Offers site persistence for e-commerce applications
- Offers a unique DNS race feature—The Cisco GSS 4480 can in real time direct content consumers to the closest data center
- Supports a Web-based graphical user interface (GUI) and DNS wizard to simplify the DNS command and control

Cisco GSS 4480 Platform Function Business Continuity and Global Service Load-Balancing Deployments

Break-Through Concept—DNS Rules

At the core of the Cisco GSS 4480 is the DNS rule, which provides centralized command and control of how the Cisco GSS 4480 globally load balances a given hosted domain, what IP address(es) is sent the client's name server (D-proxy), and what recovery method should be used if the preferred choice is unavailable. The Cisco GSS 4480 gives a DNS administrator the capability to control how Web-based applications are globally load balanced across multiple sites and under what condition clients should be directed to a backup site. This combination is perfect for DNS administrators looking for a centralized tool that gives them complete control of the DNS resolution process. The Cisco GSS 4480 provides the features needed by e-businesses looking for a real-time business continuity solution.

Using the Cisco GSS 4480 GUI, the user defines what domain will be globally load balanced. The user has the option to include the IP address of specific client named servers (D-proxy). This allows the user to “hard-wire” a single D-proxy or a group of D-proxies to a particular site. This setup is useful when the DNS administrator needs to specify a certain set of SLBs to support all the traffic coming from a specific Internet service provider (ISP) D-proxy. Next, on the same GUI screen, the DNS administrator defines what SLBs should support this domain. The DNS administrator has the option to assign multiple IP addresses (up to eight) and a Time To Live (TTL) value that will be used by the client D-proxy.

The final configuration option is the recovery method that is used if the first method fails (in the GUI these are called “DNS rules”). Each DNS rule can support up to three load-balance clauses. These clauses are used to establish the primary load-balancing method, the secondary load-balancing method, and finally the tertiary method—which could involve hard-wiring a global “sorry server.”



Cisco GSS 4480 and Cisco's Market-Leading Content Switches—Ultimate Combination for Business Resiliency Solution

The Cisco GSS 4480, in combination with a local Cisco content switch, is critical for large enterprises and service providers planning to deploy highly reliable distributed data centers. The Cisco GSS 4480 selects the best site based on the load and availability information supplied by the Cisco content switch. This leaves the Cisco content switch to select the best local server within the data center based on availability and local load. The Cisco GSS 4480 simplifies this network deployment architecture with its centralized command and control features. An example of the complete control is the fact the Cisco GSS 4480 can gracefully take a Cisco content switch out of rotation without impacting ongoing operations.

The Cisco GSS 4480 performs two major functions as part of the global site selection process:

- Takes an active role in the DNS infrastructure to connect the client to the SLB that supports the requested Web site
- Continuously monitors the load and availability of these SLBs to select the best SLB that is capable of supporting the new client

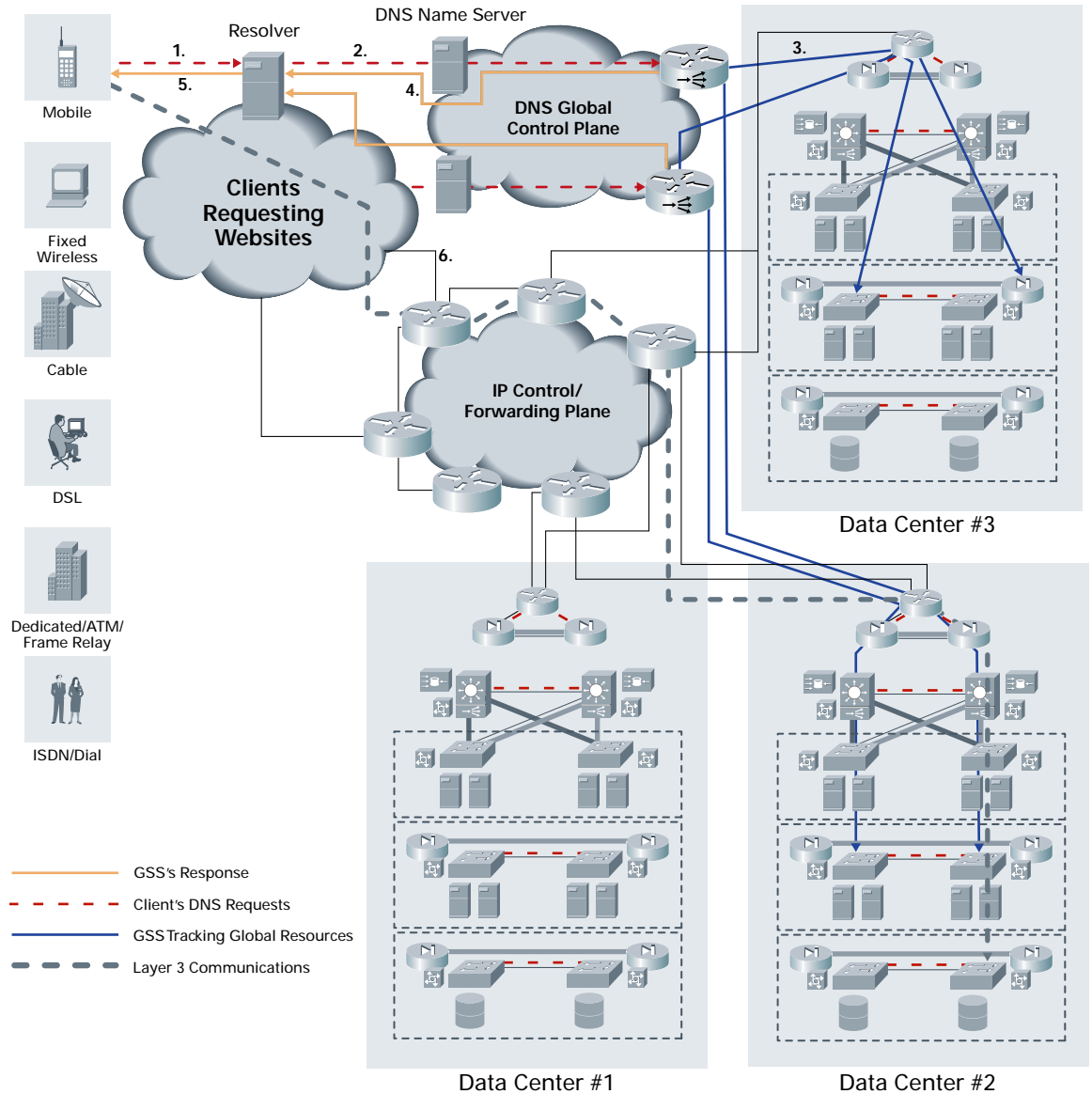
In Figure 1, the Cisco GSS 4480 offloads the Web site selection process from the DNS infrastructure. The Cisco GSS 4480 continuously monitors the load and health of up to 128 SLBs or 4000 virtual IP addresses (VIPs). These SLBs can be co-located, or located at remote and disparate data centers.

The Cisco GSS 4480 interacts with the client in the Web site selection process summarized in the following six steps:

1. A client wants to access an application at foo.com. The client types in the browser www.foo.com. This application is supported at three different data centers.
2. The request is processed by the DNS global control plane infrastructure and arrives at the Cisco GSS 4480.
3. The Cisco GSS 4480 offloads the site selection process from the DNS global control plane. The request and site selection are based on the load and health information in conjunction with customer-controlled load-balancing algorithms and the Cisco GSS 4480, in real time, selects a data center that is available and not overloaded.
4. The Cisco GSS 4480 sends the IP address of the “best” server load balancer at a specific data center, in this case the SLB at data center 2.
5. The browser processes this IP address.
6. When the hand-off for the DNS control plane is complete, the client is directed to the SLB at data center 2 by the IP control/forwarding plane.



Figure 1. Cisco GSS 4480 and Distributed Data Centers



The mechanism that is used to extract this load and health information is a specially designed keepalive (KAL). The Cisco GSS 4480 supports the following KALs:

- KAL-AP—This keepalive was developed specifically to extract both load and availability from the Cisco CSS 11000/11500 and the Cisco CSM. When this detailed query is sent to the CSS or CSM, these SLBs respond with information about a hosted domain name, or hosted VIP address of the SLB.
- Hypertext Transfer Protocol (HTTP)—A HTTP “head request” method to a given origin server sends a request to an origin server and checks for a 200 OK. If the Cisco GSS 4480 receives a 200 OK, then it directs traffic to the VIP address supporting that actual server. The user can configure the host tag, the remote host IP, and the URL (including a path).



- Internet Control Message Protocol (ICMP)—ICMP is a Layer 3 “ping” that indicates the status of a given device based on connectivity to the network. This KAL is used with any device that responds to a ping request. Pings are sent once every 45 seconds. If there is no response, the ping is sent once every five seconds up to three times. If there is still no response, the device (VIP, real sever) is considered offline.
- NS query—A simple DNS request is sent to a host (name server, mail server, or other) to receive a resolved domain name in order to prove the system is “alive.” In this case, the Cisco GSS 4480 sends a unused domain name, probing for a failure response, which proves that the DNS server is “alive.” This is used in conjunction with the name server forwarding feature.
- Content routing agent (CRA)—User Datagram Protocol (UDP) based, keepalive is used with the DNS race feature; the GSS sends keepalive requests to port 1304 to retrieve round-trip times between the GSS and agent (CSS or content engine).

Global Load-Balancing Algorithms—Complete Control of Site Selection

The Cisco GSS 4480 supports eight global load-balancing algorithms and provides administrators full flexibility in selecting the global load-balancing algorithm that matches their needs:

- Ordered list—This user-definable list specifies one or a group of IP addresses (corresponding to a VIP or the IP address of a back-end sever) that the Cisco GSS 4480 uses to respond to a DNS request for a specific domain. The Cisco GSS 4480 uses the first address in the list until it becomes unavailable or overloaded, and then it moves to the next address in the list. This process is repeated for every subsequent entry in the list.
- Static based on client’s DNS address—This algorithm is a variation of the order list but allows the administrator to map the IP address of the client’s DNS name server to an available VIP on a specific content switch. This feature is used when the administrator wants to allocate a specific community of users to a specific set of SLB or back-end srrvers.
- Round robin—This algorithm cycles through available VIPs in order. The round-robin balance method is useful when balancing requests among multiple, active data centers that are hosting identical content, for example, between SLBs at a primary and “active standby” site that serves requests.
- Weighted Round Robin—The Cisco GSS 4480 cycles through the list of VIPs that are available as requests are received, but sends requests to a favored VIP based on a user-assigned weighting value.
- Least loaded—The Cisco GSS 4480 can receive load values from either the CSS or CSM. The Cisco GSS 4480 monitors these load values to see if they exceed a threshold that is assigned by the administrator. If the load is above the specified threshold, the VIP on the CSS or CSM is deemed to be offline and unavailable to serve further requests. If the load falls below the threshold, the GSS automatically starts sending requests to the VIP.
- Source address and domain hash—When the Cisco GSS 4480 is using this balance method, the IP address of the client’s DNS proxy and the requesting client’s domain are used to create a unique “hash” value that is then “stuck” to the VIP that is chosen to serve the DNS query. This means every time that user requests that domain, the user will go to the VIP unless it is unavailable, and then the user is sent to an alternate VIP assigned by the administrator through configuration.
- DNS race—This algorithm initiates a race of A-record responses to the client’s named server. It can achieve proximity without probing.
- Drop—This algorithm silently discards requests.



The administrator can configure up to three load-balancing clauses per DNS rule. Each clause can have a different answer group assigned to it. The multiple load-balancing clauses per DNS rule allows the user to control how the Cisco GSS 4480 responds if one of the load-balancing methods fails. Using the multiple clauses per DNS rule allows the networking engineering staff to create a solution that ensures that a client always gets directed to the proper data center based on the client's unique requirements.

For example, the preferred global load-balance method may be Weighted Round Robin as defined in the first load-balancing clause. If that method fails because of the loss of a local SLB or back-end server, the Cisco GSS 4480 tries the second method, for instance, ordered list as defined in the second load-balance clause. Then if that fails, the Cisco GSS 4480 tries the third method, which may be a static entry to a global "sorry server." Users can customize the clauses based on what is most advantageous to their own environment and global load-balancing objectives.

E-Commerce

The Cisco GSS 4480 provides unique capabilities that optimize e-business. The Cisco GSS 4480 ensures reliable transactions with a "sticky" DNS feature that guarantees site persistence, ensuring "sticky" connections between users and applications during secure and nonsecure portions of the transaction. When used with a Cisco content switch, these sticky connections can eliminate dropped shopping carts and allow e-businesses to prioritize users and transactions, thereby providing premium service for high-priority customers.

Name Server Forwarding

Name server forwarding plays a vital role in providing increased flexibility for the Cisco GSS 4480. This feature is used in instances where errors in the DNS infrastructure send a request for domain that cannot be processed by the Cisco GSS 4480. The Cisco GSS 4480 does not drop the request; instead it forwards the request to a designated name server capable of responding to the request. The response of that name server is passed through the Cisco GSS 4480 such that it appears to have come from the GSS. This setup ensures that any critical DNS request will be processed and not lost.

The Cisco GSS 4480 can forward these requests using either round robin or the order list load-balancing method.

As the Cisco GSS 4480 processes these DNS requests, it can use one of the load-balancing methods to distribute these forward requests. A user can employ this feature to globally load balance e-mail servers. The following global load-balance methods can be used with the name server forwarding feature: Weighted Round Robin, ordered list, and hashing. If the Cisco GSS 4480 receives a request for an MX (mail) record, although the Cisco GSS 4480 could not respond directly, it can forward the request to one of many mail servers deployed in the enterprise or service provider network.

DNS Race—Proximity Without Probing

The Cisco GSS 4480 supports the DNS race method of proximity. This feature allows for proximity to be achieved without probing the client name server (D-proxy). It is based on a simple concept—that instantaneous proximity can be found if a device within each data center sends an A-record (IP address), at the exact same time, to the client's named server. Whichever A-record is received first is the most proximate.



The DNS race method of DNS resolution is initiated by the Cisco GSS 4480 and is designed to load balance up to 20 sites. In order for the Cisco GSS 4480 to initiate a race, it needs to establish two pieces of information per CRA:

- The delay between the Cisco GSS 4480 and each of the CRAs in each data center—With this data, the Cisco GSS 4480 computes how much time to delay the race from each data center so each CRA will start the race simultaneously.
- The “aliveness” of the CRAs—With this data, the Cisco GSS 4480 knows not to forward a request to any CRAs that are not responding.

The Cisco GSS 4480 gathers this information by sending keepalive messages at predetermined intervals. This data, along with the IP addresses of the CRA, is used to request the start of the race. When the Cisco GSS 4480 receives a DNS request, a race request is sent to each CRA at a predetermined time and the race is initiated from each data center. The first A-record received by the client’s D-proxy is the winner and is the most proximate. The DNS race method on the Cisco GS 4480 is exactly what a DNS administrator is looking for as a way to reduce the time involved in the DNS resolution process.

Scalability

Highly scalable, the Cisco GSS 4480 is designed to meet the most demanding environments. Cisco GSS 4480 support information is given in Table 1.

Table 1 Cisco GSS 4480 Performance and Scalability Metrics

DNS requests per second	4000
DNS rules	4000
Name server forwarding requests per second or DNS races per second	1500
Active SLBs	128
Hosted domains (maximum 1000 per SLB), 128 characters maximum per domain	2000
Hosted domain lists (500 members maximum per list)	250
VIPs (no more than 500 with active Internet Control Message Protocol [ICMP] keepalive)	4000
Name server addresses for name server forwarding (30 maximum per answer group)	100
CRA devices (20 maximum per race and answer group)	200
Source IP addresses configurable for DNS rules	500
Source address lists (30 members maximum per list)	60
Answer groups (100 members maximum per group)	500



Manage Through Simple GUIs

For configuration and control, administrators have a Cisco IOS[®] Software-like command-line interface (CLI) and an intuitive, embedded GUI. The CLI is used for network startup, and the GUI is used for the configuration of all global load-balancing parameters and is supported directly by the Cisco GSS 4480. Users have a choice of using either an expert mode on the GUI or a “wizard” to configure all global load-balancing parameters at a single, centralized location. Configuration parameters are shared with up to eight Cisco GSS 4480s in a single deployment. Access to the CLI is via Telnet or Secure Shell Protocol (SSH); the GUI is accessed via Secure Sockets Layer (SSL). Network managers can restrict console, File Transfer Protocol (FTP), Simple Network Management Protocol (SNMP), Telnet, and Web management access.

Monitoring the status of the Cisco GSS 4480 can be done via GUI, SNMP, CLI, and log files. The information that can be accessed from the different methods includes:

- Displaying of configuration, software information, hardware information, system resources, user information, and current logins
- Booting configuration
- Configuring IP interfaces, IP routes, host name
- The ability to see the operation status (active/inactive) of locally configured resources, with filtering by domain name or other parameters (GUI only)

Configuration State and File Management

- Show running configuration for all CLI-configured options
- Show startup configuration
- Archive and restore running configuration
- Archive and restore startup configuration
- Display archived configuration
- Copy running configuration to disk, FTP, or Trivial File Transfer Protocol (TFTP) server
- Clear running configuration
- Create startup configuration, including from text editor
- Copy startup-config to running-config

Logging

- Show logs, including real-time log
- Archive and restore log files



Cisco GSS 4480 Features: Quick Look

Table 2 gives features of the Cisco GSS 4480.

Table 2 Cisco GSS 4480 Features

Cisco GSS 4480	
Ports	Two 10/100 Fast Ethernet
DNS requests per second	4000, depending on configuration (~ 345 million DNS requests per day per Cisco GSS 4480; an entire system is capable of 2.7 billion DNS requests per day)
Devices supported	CSSs—Load and availability CSMs—Load and availability Cisco Local Director—Availability using HTTP head KAL Cisco IOS SLB—Availability using HTTP head KAL DNS name servers—Availability using name server request KAL Origin servers—Availability using HTTP head KAL Content engines—Availability using ICMP ping KAL CRAs—Availability using CRA KAL
Network management	Console port—CLI Access to system via Telnet Secure copy (SCP) or FTP GUI—Secure HTTP (HTTPS) for Internet Explorer and Netscape Navigator
Network management Management Information Bases (MIBs)	Read-only monitoring of network and device status, including RFC 1213 (MIB-II) and RFC 1514 (HOST-RESOURCES-MIB)
Physical	One-rack unit size chassis Network management serial port 1 GB of RAM 600-MHz PIII CPU
Storage	One 36-GB hard drive
Power	Integrated AC power (auto-sense 110V/60 Hz)



Ordering Information

The Cisco GSS 4480 supports only AC power and there are no options. Orderable part numbers are given in Table 3.

Table 3 Part Numbers for Cisco GSS 4480

Product number	Description
Cisco GSS 4480-K9	Global site selector
SF-GSS-V1.0-K9	Global site selector software

Environmental Specifications

Table 4 gives specifications of the Cisco GSS 4480.

Table 4 Specifications of Cisco GSS 4480

Rack units	One rack unit
Ports	Two 10/100 Fast Ethernet auto-sensing, one console port
Storage	One 36-GB hard drive
Software image	SF-GSS-V1.0-K9
Processor	600-MHz PIII CPU
Mechanical and environmental	
Dimensions	H x W x D): 1.72 x 17.5 x 14.13 in. (43.7 x 444.5 x 358.9 cm)
Weight	12.5 lb (5.7 kg)
Operational temperature	32 to 113°F (0 to 45°C)
Storage temperature	-13 to 158°F (-25 to 70°C)
Relative humidity	95%, nonoperational
Certification	
Safety	UL 1950; CAS 950; EN60950: GS Mark
EMC	FCC Part 15 Class A; EN55022 Class A; EN50082-1; VCCI Class 1
Factory approvals	UL; TUV; BABT 340

**Corporate Headquarters**

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters

Cisco Systems Europe
11, Rue Camille Desmoulins
92782 Issy-les-Moulineaux
Cedex 9
France
www-europe.cisco.com
Tel: 33 1 58 04 60 00
Fax: 33 1 58 04 61 00

Americas Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters

Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: 65 317 7777
Fax: 65 317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the Cisco Web site at www.cisco.com/go/offices

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia
Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland
Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland
Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Copyright © 2002, Cisco Systems, Inc. All rights reserved. CCIP, the Cisco Powered Network mark, the Cisco Systems Verified logo, Cisco Unity, Follow Me Browsing, FormShare, iQ Breakthrough, iQ Expertise, iQ FastTrack, the iQ logo, iQ Net Readiness Scorecard, Networking Academy, ScriptShare, SMARTnet, TransPath, and Voice LAN are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn, Discover All That's Possible, The Fastest Way to Increase Your Internet Quotient, and iQuick Study are service marks of Cisco Systems, Inc.; and Aironet, ASIST, BPX, Catalyst, CCDA, CCDP, CCIE, CCNA, CCNP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, the Cisco IOS logo, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Empowering the Internet Generation, Enterprise/Solver, EtherChannel, EtherSwitch, Fast Step, GigaStack, Internet Quotient, IOS, IP/TV, LightStream, MGX, MICA, the Networkers logo, Network Registrar, Packet, PIX, Post-Routing, Pre-Routing, RateMUX, Registrar, SlideCast, StrataView Plus, Stratm, SwitchProbe, TeleRouter, and VCO are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site 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. (0208R)

Printed in the USA