

Cisco Catalyst 4500 E-Series High Availability

Introduction

High availability is a critical requirement of most networks. Minimizing Ethernet switch downtime maximizes productivity for hosts and other network devices connected to it. The Cisco® Catalyst® 4500 E-Series provides several features to minimize planned and unplanned outages.

High availability in the Cisco Catalyst 4500 E-Series is achieved through these features:

- In-Service Software Upgrade (ISSU)*
- Stateful Switchover (SSO)*
- Nonstop Forwarding (NSF)*
- Supervisor Engine Uplink Redundancy
- Gateway Load Balancing Protocol (GLBP)
- EtherChannel®
- Power Supply Redundancy

* Available Q1CY '08

In-Service Software Upgrade (ISSU)

What Is ISSU?

ISSU, available on the Cisco Catalyst 4500 E-Series, allows customers to virtually eliminate planned outages for full feature software upgrades. It provides the means to upgrade or, if needed, downgrade the Cisco IOS® Software in a redundant Cisco Catalyst 4500 E-Series system without incurring a service outage. ISSU adds additional functionality to the Cisco Catalyst 4500 E-Series high-availability capabilities provided by Stateful Switchover (SSO) and Nonstop Forwarding (NSF), discussed elsewhere in this paper. ISSU is a user-initiated and user-controlled process. It is carried out through a set of executive-level CLI commands issued in a specific order to upgrade or downgrade a Cisco IOS Software image running on a Cisco Catalyst 4500 E-Series dual supervisor engine configuration. ISSU differs from other “hitless” software upgrades in that it provides the ability to do a hitless “full feature” upgrade rather than just a system patch.

How Does ISSU Work?

ISSU should be thought of as a process allowing customers to upgrade or if needed downgrade a Cisco IOS Software image running on a Cisco Catalyst 4500 E-Series system configured for SSO/NSF, from a lower version to a higher version or vice versa. This process moves a Cisco Catalyst 4500 E-Series from one version of an SSO/NSF-capable Cisco IOS Software image to another version of an SSO/NSF-capable Cisco IOS Software image with minimized downtime, degradation of service, and loss of packets. In order to perform ISSU with the Cisco Catalyst 4500 E-Series, you must first have the following:

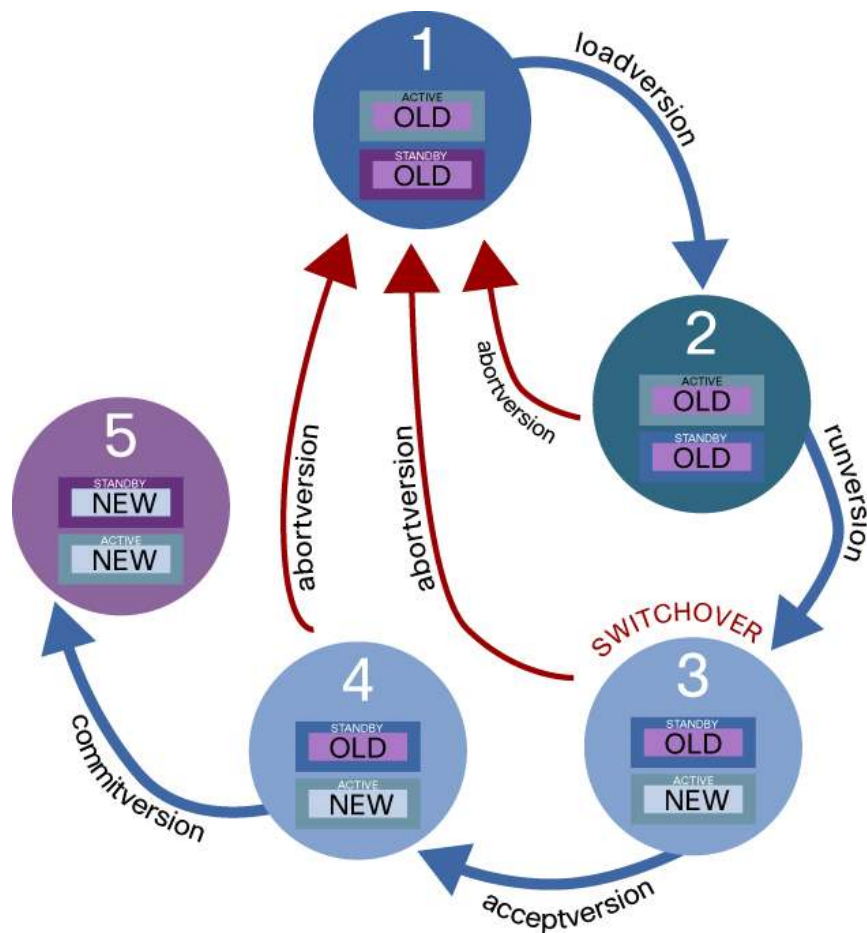
- A redundant Cisco Catalyst 4500 chassis (4507R-E or 4510R-E)
- Previous implementation of Cisco NSF/SSO

The ISSU itself is a user-initiated and user-controlled process implemented through a set of exec-level CLI commands issued in a specific order. In essence, the procedure can be simply described as follows:

1. Reset the standby supervisor engine with the new software.
2. Switch over to the standby supervisor engine with the new software; making it the active supervisor engine.
3. Reset the new standby supervisor engine (the original active supervisor engine) with the new software.

The ISSU process and the commands used to execute each step of the process are illustrated in Figure 1.

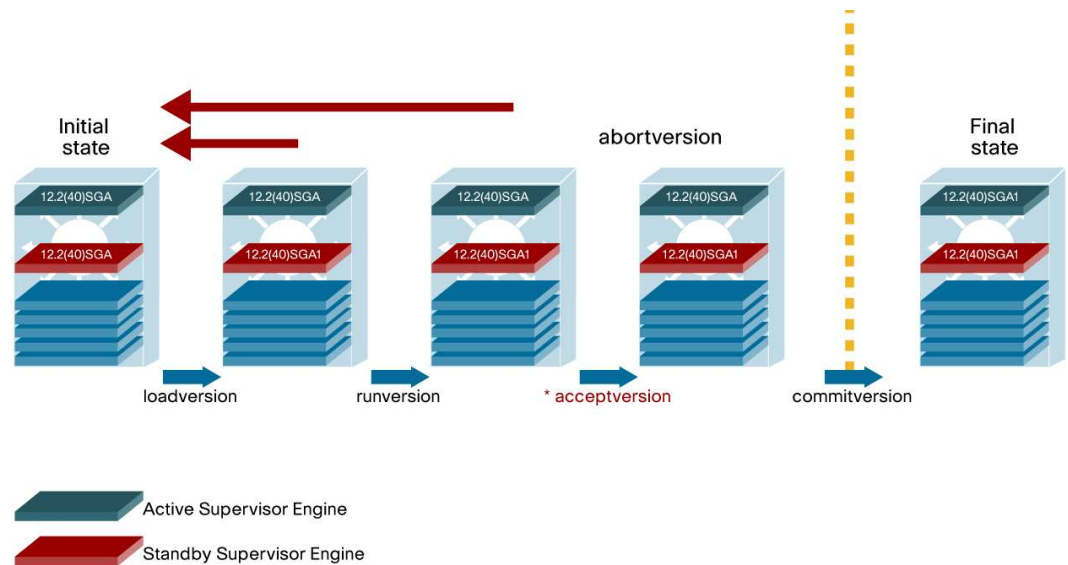
Figure 1. ISSU Process



ISSU Prerequisites

Before one can perform an ISSU, there are a few prerequisites to be verified. The following list explains what is initially required.

- Must be using a redundant Cisco Catalyst 4500 E-Series switch with symmetric hardware (that is, dual Cisco Catalyst 4500 Series Supervisor Engines of the same model).
- Both new and old Cisco IOS Software images must be preloaded to the file system on both supervisor engines.



The ISSU process is started by entering the **issu loadversion** command on the active supervisor engine. This command directs the active supervisor engine to begin the ISSU process. The active supervisor engine, through inter-supervisor engine communications, checks that the requested image has been downloaded into both the active and standby supervisor engines' file systems. If the required images are not present, the command is rejected, and a warning message is displayed on the system console. If the **issu loadversion** command is successful, the switch transitions into the "Load Version" ISSU state. The standby supervisor engine will reset and boot with the new version of the Cisco IOS Software image loaded into the file system. The following actions take place when the command is implemented:

1. The standby supervisor engine is reset.
2. The standby supervisor engine is booted with the new Cisco IOS Software image.
3. If both Cisco IOS Software images are declared as compatible, the standby supervisor engine moves into SSO mode and is fully stateful. Compatibility allows the process to proceed from step 5 below.
4. If both Cisco IOS Software images are incompatible, the system moves into route processor redundancy (RPR) mode, and the ISSU process is terminated with an appropriate message to the user. Images are declared incompatible when "required" clients or applications are not interoperable between two Cisco IOS Software releases.
5. Standby "B" reaches the standby HOT state.
6. The user has an option to abort the ISSU process by issuing the **issu abortversion** command.

The second step of the ISSU process is to perform the **issu runversion** CLI. The user can issue the **issu runversion** command when:

1. The ISSU state is "Load Version"; this can be verified with the **show issu state detail** command.
2. The standby supervisor engine is running the new version of the software.
3. The standby supervisor engine has moved into the "Standby Hot" state.

The following actions take place when the **issu runversion** command is executed:

1. A switchover occurs; that is, the standby (B) becomes the new active, and the old active (A) is rebooted with the new software and comes up as a standby.
2. A timer called "Rollback Timer" is started with a default or previously configured value.
3. Move both supervisor engines to "Run Version" state.
4. If the command **issu acceptversion** is not issued before the "Rollback timer" expires, then the entire ISSU process is aborted using the automatic rollback.
5. When the **issu acceptversion** command is issued the rollback timer is stopped. This command must be issued from the console connection of the new active supervisor.
6. The user has an option to abort the ISSU process by issuing the **issu abortversion** command.

As soon as the **issu runversion** command is issued and both supervisor engines initiate the transition into the "Run Version" state, a timer called the "Rollback Timer" is started. The purpose of this timer is to provide the user with a window to help ensure that the newly active supervisor engine is reachable both from console as well as the network. Additionally, the user can use this window of time to make sure that certain critical features and functionality are working. If the user feels that a longer window of time is needed to help ensure such functionality, the value of the timer can be extended. Once the user is satisfied that the new image is performing correctly, the user either issues an **issu acceptversion** command to proceed with new image. Alternatively, if the new image is not performing correctly the user may issue the **issu abortversion** command to go back to the previous version. Both commands stop the rollback timer. The value of the rollback timer is a configurable parameter beginning from zero (meaning disable rollback) to 2 hours; default value is 45 minutes. To change the default value of the rollback timer, one needs to configure the timer settings before starting the ISSU process. If the rollback timer expires, the ISSU process is terminated.

The following actions take place in order to prepare for and to implement rollback:

1. The ISSU state for both supervisor engines is set to "INIT."
2. The boot parameters of both supervisor engines are updated to refer to the old image version.
3. An event is logged on the active supervisor engine indicating that the rollback process has begun.
4. The switchover is done similar to the one mentioned in **issu runversion** with the standby supervisor engine becoming the active.
5. A Simple Network Management Protocol (SNMP) trap is generated reporting the failure reason and previous and next states.
6. Switchovers are enabled.

The following actions take place when the **issu acceptversion** command is executed:

1. The "Rollback Timer" is terminated. This means that the rollback timer is not looked at anymore. Therefore, the system will continue to run in this state. The **issu acceptversion** command halts the rollback timer and helps ensure the ISSU process is not automatically aborted.

Committing the new image to the standby supervisor engine is the last stage of the ISSU procedure. When committing the new Cisco IOS Software image on the active supervisor engine, the standby supervisor engine resets and reloads with the new image. Issuing the **issu**

commitversion command initiates the commit process. This command resets the standby supervisor engine and boots it with a new version of the software (the same now running on the active supervisor engine). This concludes the ISSU process, and the new version of software is operating on both supervisor engines. Since this is the conclusion of the ISSU process, the system cannot be reverted back to the previous version of the software from this point onward as a part of this upgrade cycle. However, if for any reason users wish to go back to the previous version of the software, they can do so by starting a new downgrade process.

The following actions take place when the **issu commitversion** command is executed:

1. The standby supervisor engine (A) is reset and booted with the new version of Cisco IOS Software image.
2. The standby supervisor engine (A) moves into the "Standby Hot" state in SSO mode and is fully stateful for all clients/applications that are compatible with ISSU.
3. Both supervisor engines are moved into "Final State," which is the same as "Initial State."
4. Administrators can initiate switchovers from this point onward.

One can abort the ISSU process at any stage (prior to issuing the **issu commitversion** command) by issuing the **issu abortversion** command. The ISSU process also aborts on its own if the software detects a failure. If a user aborts the process after issuing the **issu loadversion** command, then the standby supervisor engine is reset and reloaded with the original software. If the process is aborted after a user enters either the **issu runversion** or **issu acceptversion** command, then a second switchover is performed to the new standby supervisor engine that is still running the original software version. The supervisor engine that had been running the new software is reset and reloaded with the original software version. The command is accepted only in "Load Version" or "Run Version" state. In "Load Version" state, the active supervisor engine is running an old image, and the standby supervisor engine is running a new image.

Stateful Switchover (SSO)

What Is SSO?

ISSU depends on the facilities provided by SSO. Cisco Catalyst 4500 E-Series Switches allow a redundant supervisor engine to quickly take over operation of the switch if the active supervisor engine fails. Supervisor engine redundancy is enabled by running the redundant supervisor engine in SSO operating mode. With supervisor engine redundancy enabled, if the active supervisor engine fails, ISSU commands are issued, or a manual switchover is performed, the redundant supervisor engine becomes the active supervisor engine. The redundant supervisor engine is automatically initialized with the startup configuration of the active supervisor engine. This shortens the switchover time from 30 seconds or longer in RPR mode to less than 200 milli-seconds in SSO mode.

When a redundant supervisor engine runs in SSO mode, it starts up in a fully initialized state and synchronizes with the persistent configuration and the running configuration of the active supervisor engine. It subsequently maintains the state of SSO client protocols. All changes in hardware and software states for features that support SSO are kept in sync. Consequently, it offers zero interruption to Layer 2 sessions in a redundant supervisor engine configuration.

Because the redundant supervisor engine recognizes the hardware link status of every link, ports that were active before the switchover remain active. This includes the uplink ports. However, because uplink ports are physically on the supervisor engine, they are disconnected if the

supervisor engine is removed. If the active supervisor engine fails, the redundant supervisor engine becomes active. This newly active supervisor engine uses Layer 2 switching information that exists to continue forwarding traffic. Unless NSF is configured, Layer 3 forwarding is delayed until the routing tables have been repopulated in the newly active supervisor engine.

What Is Synchronized?

SSO supports these Layer 2 features. The state of these features is preserved between both the active and redundant supervisor engines:

- 802.3
- 802.3u
- 802.3x (Flow Control)
- 802.3ab (Gigabit Ethernet)
- 802.3z (Gigabit Ethernet including coarse wavelength division multiplexing [CWDM])
- 802.3ad (Link Aggregation Control Protocol [LACP])
- 802.1p (Layer 2 quality of service [QoS])
- 802.1q
- 802.1X (authentication)
- 802.1D (Spanning Tree Protocol)
- 802.3af (inline power)
- Port Aggregation Protocol (PAgP)
- Virtual Trunk Protocol (VTP)
- Dynamic Address Resolution Protocol (ARP) Inspection
- DHCP snooping
- IP source guard
- Internet Group Management Protocol (IGMP) snooping (versions 1 and 2)
- Distributed Diagnostics and Service Network (DDSN) Transfer Protocol (DTP)
- Multiple Spanning Tree (MST)
- Per-VLAN Spanning Tree (PVST+)
- Rapid PVST
- PortFast/UplinkFast/BackboneFast
- Bridge Protocol Data Unit (Bpdu) guard and filtering
- Voice VLAN
- Port security
- Unicast MAC filtering
- Access Control List (ACL), VLAN Access Control List (VACL), Port Access Control List (PACL), Receive Access Control List (RACL)
- QoS (Dynamic Buffer Limiting [DBL])
- Multicast storm control/broadcast storm control

What Is Not Synchronized?

These are not synchronized between the active and standby supervisor engines:

- All Layer 3 protocols on Cisco Catalyst 4500 E-Series Switches (Switch Virtual Interfaces)
Cisco NSF works with SSO to minimize the amount of time that a Layer 3 network is unavailable after a supervisor engine switchover by continuing to forward IP packets. The reconvergence of Layer 3 routing protocols (Border Gateway Protocol [BGP], Enhanced Interior Gateway Routing Protocol [EIGRP], Open Shortest Path First [OSPF] v2, and Intermediate System-to-Intermediate System [IS-IS]) is transparent to the user and occurs automatically in the background. The routing protocols recover routing information from neighbor devices and rebuild the Cisco Express Forwarding table. SSO is compatible with this list of features. However, the protocol database for these features is not synchronized between the redundant and active supervisor engines.
- 802.1Q tunneling with Layer 2 Protocol Tunneling
- Baby giants
- Jumbo frame support
- Cisco Discovery Protocol
- Flood blocking
- Unidirectional Link Detection Protocol (UDLD)
- Switched Port Analyzer (SPAN)/Remote Switch Port Analyzer (RSPAN)

Configuration

Issue these commands in order to configure the redundancy in SSO mode:

```
Switch#configure terminal  
Switch(config)#redundancy  
Switch(config-red)#mode sso
```

Verification of SSO

These commands verify supervisor engine redundancy on the Cisco Catalyst 4500 E-Series Switches:

- **show module:** This verifies whether the redundant supervisor engine module exists and is in standby mode.
- **show redundancy:** This verifies the redundancy facility information.

Manual Switchover Commands

Manual switchover commands can be used to perform move control of the switch from the active supervisor engine to the redundant supervisor engine. The redundancy force-switchover command launches the switchover only if the state of the redundant supervisor engine is Standby Hot. If the state is not Standby Hot, the command does not process. Issue the redundancy force-switchover command, rather than the reload command, to initiate a switchover. The redundancy force-switchover command first checks that the redundant supervisor engine is in the correct state. If you issue the reload command and the status is not Standby Hot, the reload command only resets the current supervisor engine.

Nonstop Forwarding (NSF)

What Is NSF?

NSF works with SSO to minimize the amount of time the network is unavailable following a supervisor engine switchover.

NSF provides these benefits:

- **Improved network availability:** NSF continues forwarding network traffic and application state information so that user session information is maintained after a switchover.
- **Prevents routing flaps:** NSF continues forwarding routing traffic while reestablishing routing relationships
- **Improves network stability:** Network stability is improved by avoiding routing flaps.
- **Neighboring routers do not detect a link flap:** Because the interfaces remain up during a switchover, neighboring routers do not detect a link flap.
- **Maintains user sessions** established prior to the switchover.

NSF uses capabilities of Layer 3 routing protocols and Cisco Express Forwarding to prevent disruption of traffic forwarding. The BGP, OSPF, and EIGRP routing protocols have been enhanced with NSF capability and awareness, which means that routers running these protocols can detect a switchover and take the necessary actions to continue forwarding network traffic and to recover route information from the peer devices. The IS-IS protocol can be configured to use state information that has been synchronized between the active and the redundant supervisor engine to recover route information following a switchover instead of information received from peer devices.

The routing protocols run only on the active supervisor engine, and they receive routing updates from their neighbor routers. Routing protocols do not run on the standby supervisor engine. Following a switchover, the routing protocols request that the NSF-aware neighbor devices send state information to help rebuild the routing tables. Alternately, the IS-IS protocol can be configured to synchronize state information from the active to the redundant supervisor engine to help rebuild the routing table on the Cisco Catalyst 4500 E-Series switch in environments where neighbor devices are not NSF-aware. Each protocol depends on Cisco Express Forwarding to continue sending packets during switchover while the routing protocols rebuild the Routing Information Base (RIB) tables. After the routing protocols have converged, Cisco Express Forwarding updates the FIB table and removes stale route entries. Cisco Express Forwarding maintains the FIB and uses the FIB information that was current at the time of the switchover to continue sending packets during a switchover. This feature reduces traffic interruption during the switchover. During normal NSF operation, Cisco Express Forwarding on the active supervisor engine synchronizes its current FIB and adjacency databases with the FIB and adjacency databases on the redundant supervisor engine. Upon switchover of the active supervisor engine, the redundant supervisor engine initially has FIB and adjacency databases that are mirror images of those that were current on the active supervisor engine. As the routing protocols start to repopulate the RIB on a prefix-by-prefix basis, the updates cause prefix-by-prefix updates to Cisco Express Forwarding, which it uses to update the FIB and adjacency databases. Existing and new entries receive the new version ("epoch") number, indicating that they have been refreshed. The forwarding information is updated on the forwarding engine during convergence. The supervisor engine signals when the RIB has converged. The software removes all FIB and adjacency entries that have an epoch older than the current switchover epoch. The FIB now represents the newest routing protocol forwarding information.

BGP NSF Operation

When a Cisco Catalyst 4500 E-Series switch with NSF begins a BGP session with a BGP peer, it sends an OPEN message to the peer. Included in the message is a statement that the switch has “graceful” restart capability. Graceful restart is the mechanism by which BGP routing peers avoid a routing flap following a switchover. If the BGP peer has received this capability, it is aware that the Catalyst 4500 Series switch sending the message is NSF-capable. Both the NSF-capable switch and its BGP peers need to exchange the graceful restart capability in their OPEN messages at the time of session establishment. If both the peers do not exchange the graceful restart capability, the session will not be capable of a graceful restart.

If the BGP session is lost during the supervisor engine switchover, the NSF-aware BGP peer marks all the routes associated with the NSF-capable Cisco Catalyst 4500 E-Series switch as stale; however, it continues to use these routes to make forwarding decisions for a set period of time. This functionality prevents packets from being lost while the newly active supervisor engine is waiting for convergence of the routing information with the BGP peers.

After a supervisor engine switchover occurs, the NSF-capable Cisco Catalyst 4500 E-Series switch reestablishes the session with the BGP peer. In establishing the new session, it sends a new graceful restart message that identifies the itself as having restarted.

At this point, the routing information is exchanged between the two BGP peers. After this exchange is complete, the NSF-capable Cisco Catalyst 4500 E-Series switch uses the routing information to update the RIB and the FIB with the new forwarding information. The NSF-aware Cisco Catalyst 4500 E-Series switch uses the network information to remove stale routes from its BGP table; the BGP protocol then is fully converged.

If a BGP peer does not support the graceful restart capability, it ignores the graceful restart capability in an OPEN message but establishes a BGP session with the NSF-capable Cisco Catalyst 4500 E-Series switch. This function allows interoperability with non-NSF-aware BGP peers (and without NSF functionality), but the BGP session with non-NSF-aware BGP peers is not capable of a graceful restart.

BGP support in NSF requires that neighbor networking devices be NSF-aware; that is, the devices must have the graceful restart capability and advertise that capability in their OPEN message during session establishment. If an NSF-capable Cisco Catalyst 4500 E-Series switch discovers that a particular BGP neighbor does not have graceful restart capability, it does not establish an NSF-capable session with that neighbor. All other neighbors that have graceful restart capability continue to have NSF-capable sessions with this NSF-capable Cisco Catalyst 4500 E-Series switch.

NSF support for BGP is configured using the **bgp graceful-restart** command in router configuration mode.

OSPF NSF Operation

When an OSPF NSF-capable Cisco Catalyst 4500 E-Series switch performs a supervisor engine switchover, it must perform the following tasks in order to resynchronize its link state database with its OSPF neighbors:

- Relearn the available OSPF neighbors on the network without causing a reset of the neighbor relationship
- Reacquire the contents of the link state database for the network

As quickly as possible after a supervisor engine switchover, the NSF-capable Cisco Catalyst 4500 E-Series switch sends an OSPF NSF signal to neighboring NSF-aware devices. Neighbor networking devices recognize this signal as an indicator that the neighbor relationship with this router should not be reset. As the NSF-capable Cisco Catalyst 4500 E-Series switch receives signals from other routers on the network, it can begin to rebuild its neighbor list.

After neighbor relationships are reestablished, the NSF-capable Cisco Catalyst 4500 E-Series switch begins to resynchronize its database with all of its NSF-aware neighbors. At this point, the routing information is exchanged between the OSPF neighbors. Once this exchange is complete, the NSF-capable Cisco Catalyst 4500 E-Series switch uses the routing information to remove stale routes, update the RIB, and update the FIB with the new forwarding information. The OSPF protocols are then fully converged.

OSPF support in NSF requires that all neighbor networking devices be NSF-aware. If an NSF-capable Cisco Catalyst 4500 E-Series switch discovers that it has non-NSF-aware neighbors on a particular network segment, it disables NSF capabilities for that segment. Other network segments composed entirely of NSF-capable or NSF-aware routers continue to provide NSF capabilities.

NSF support for OSPF is configured using the **nsf** command in router configuration mode.

IS-IS NSF Operation

When an IS-IS NSF-capable Cisco Catalyst 4500 E-Series switch performs a supervisor engine switchover, it must perform the following tasks in order to resynchronize its link state database with its IS-IS neighbors:

- Relearn the available IS-IS neighbors on the network without causing a reset of the neighbor relationship
- Reacquire the contents of the link state database for the network

The IS-IS NSF feature offers two options when you configure NSF:

- Internet Engineering Task Force (IETF) IS-IS
- Cisco IS-IS

If neighbor routers on a network segment are running a software version that supports the IETF Internet draft for router restartability, they assist an IETF NSF router that is restarting. With IETF, neighbor routers provide adjacency and link-state information to help rebuild the routing information following a switchover. A benefit of IETF IS-IS configuration is operation between peer devices based on a proposed standard.

As quickly as possible after a supervisor engine switchover, the NSF-capable Cisco Catalyst 4500 E-Series switch sends IS-IS NSF restart requests to neighboring NSF-aware devices using the IETF IS-IS configuration. Neighbor networking devices recognize this restart request as an indicator that the neighbor relationship with this Cisco Catalyst 4500 E-Series switch should not be reset, but that they should initiate database resynchronization with it. As the restarting Cisco Catalyst 4500 E-Series switch receives restart request responses from routers on the network, it can begin to rebuild its neighbor list.

After this exchange is complete, the NSF-capable Cisco Catalyst 4500 E-Series switch uses the link-state information to remove stale routes, update the RIB, and update the FIB with the new forwarding information; IS-IS is then fully converged.

The switchover from one supervisor engine to the other happens within 200 milliseconds. IS-IS reestablishes its routing table and resynchronizes with the network within a few additional seconds. At this point, IS-IS waits for a specified interval before it attempts a second NSF restart. During this time, the new redundant supervisor engine boots up and synchronizes its configuration with the active supervisor engine. The IS-IS NSF operation waits for a specified interval to make sure that connections are stable before attempting another restart of IS-IS NSF. This functionality prevents IS-IS from attempting back-to-back NSF restarts with stale information.

Using the Cisco configuration option, full adjacency and LSP information is saved, or *checkpointed*, to the redundant supervisor engine. Following a switchover, the newly active supervisor engine maintains its adjacencies using the checkpointed data and can quickly rebuild its routing tables.

NSF support for IS-IS is configured using the **nsf** command in router configuration mode.

EIGRP NSF Operation

When an EIGRP NSF-capable Cisco Catalyst 4500 E-Series switch initially reboots after an NSF restart, it has no neighbor, and its topology table is empty. The switch is notified by the redundant (now active) supervisor engine when it needs to bring up the interfaces, reacquire neighbors, and rebuild the topology and routing tables. The restarting Cisco Catalyst 4500 E-Series switch and its peers must accomplish these tasks without interrupting the data traffic directed toward the restarting switch. EIGRP peer routers maintain the routes learned from the restarting router and continue forwarding traffic through the NSF restart process.

To prevent an adjacency reset by the neighbors, the restarting switch uses the Restart (RS) bit in the EIGRP packet header to indicate a restart. The RS bit is set in the hello packets and in the initial INIT update packets during the NSF restart period. The RS bit in the hello packets allows the neighbors to be quickly notified of the NSF restart. Without seeing the RS bit, the neighbor can only detect an adjacency reset by receiving an INIT update or by the expiration of the hello hold timer. Without the RS bit, a neighbor does not know if the adjacency reset should be handled using NSF or the normal startup method.

When the neighbor receives the restart indication, by receiving either the hello packet or the INIT packet, it recognizes the restarting peer in its peer list and maintains the adjacency with the restarting router. The neighbor then sends its topology table to the restarting router with the RS bit set in the first update packet, indicating that it is NSF-aware and is helping out the restarting switch. The neighbor does not set the RS bit in its hello packets, unless it is also an NSF restarting neighbor.

If at least one of the peer routers is NSF-aware, the restarting Cisco Catalyst 4500 E-Series switch would then receive updates and rebuild its database. The restarting switch must then find out if it had converged so that it can notify the routing information base (RIB). Each NSF-aware router is required to send an end of table (EOT) marker in the last update packet to indicate the end of the table content. The restarting switch knows it has converged when it receives the EOT marker. The restarting Cisco Catalyst 4500 E-Series switch can then begin sending updates.

An NSF-aware peer would know when the restarting Cisco Catalyst 4500 E-Series switch had converged when it receives an EOT indication from it. The peer then scans its topology table to search for the routes with the restarted neighbor as the source. The peer compares the route timestamp with the restart event timestamp to determine if the route is still available. The peer then

goes active to find alternate paths for the routes that are no longer available through the restarted switch.

When the restarting switch has received all EOT indications from its neighbors or when the NSF converge timer expires, EIGRP notifies the RIB of convergence. EIGRP waits for the RIB convergence signal and then floods its topology table to all awaiting NSF-aware peers.

NSF support for EIGRP is configured using the **nsf** command in router configuration mode.

Supervisor Engine Uplink Redundancy

The Supervisor Engine 6-E includes two 10 Gigabit Ethernet uplinks. The Cisco Catalyst 4500 E-Series is designed such that the uplink ports on both the active and backup supervisor engines can be active and forwarding traffic at the same time. This capability is enabled from software Cisco IOS Software Release 12.2(40)SGA.

The uplinks on the Supervisor Engine 6-E can operate in several modes. When both ports on each supervisor engine are forwarding traffic, the configuration is referred to as 2+2 redundancy. In this so-called 2+2 configuration, the uplinks are oversubscribed 2:1. When only one port on each of the two supervisor engines carries traffic, a configuration referred to as 1+1 redundancy, each port forwards traffic at line rate.

All the active links (on either active or redundant Supervisor engine) will remain active even during a switchover scenario (SSO/NFS/ISSU) as described earlier.

The Supervisor Engine 6-E also supports the TwinGig Converter module in one or both of the two 10 Gigabit Ethernet uplink ports. The TwinGig Converter module provides two Gigabit Ethernet Small Form-Factor Pluggable (SFP) ports. Using this module, the Cisco Catalyst 4500 E-Series supports 4+4 uplink redundancy wherein all four 1 Gigabit uplinks on each of two supervisor engines carry traffic simultaneously. When operating in this mode all four ports carry traffic at line rate.

Gateway Load Balancing Protocol

The GLBP feature provides automatic router backup for IP hosts configured with a single default gateway on an IEEE 802.3 LAN. Multiple first hop routers on the LAN combine to offer a single virtual first hop IP router while sharing the IP packet forwarding load. Other routers on the LAN may act as redundant GLBP routers that will become active if any of the existing forwarding routers fails.

GLBP performs a similar, but not identical, function for the user as the Hot Standby Routing Protocol (HSRP) and the Virtual Router Redundancy Protocol (VRRP). HSRP and VRRP protocols allow multiple routers to participate in a virtual router group configured with a virtual IP address. One member is elected to be the active router to forward packets sent to the virtual IP address for the group. The other routers in the group are redundant until the active router fails. These standby routers have unused bandwidth that the protocol is not using. Although multiple virtual router groups can be configured for the same set of routers, the hosts must be configured for different default gateways, which results in an extra administrative burden. GLBP provides load balancing over multiple routers (gateways) using a single virtual IP address and multiple virtual MAC addresses. Each host is configured with the same virtual IP address, and all routers in the virtual router group participate in forwarding packets.

Members of a GLBP group elect one gateway to be the active virtual gateway (AVG) for that group. Other group members provide backup for the AVG in the event that the AVG becomes unavailable. The AVG assigns a virtual MAC address to each member of the GLBP group. Each gateway assumes responsibility for forwarding packets sent to the virtual MAC address assigned to it by the AVG. These gateways are known as active virtual forwarders (AVFs) for their virtual MAC address.

The AVG is responsible for answering Address Resolution Protocol (ARP) requests for the virtual IP address. Load sharing is achieved through the AVG sending ARP replies with different virtual MAC addresses.

A GLBP group allows up to four virtual MAC addresses per group. The AVG is responsible for assigning the virtual MAC addresses to each member of the group. Other group members request a virtual MAC address after they discover the AVG through hello messages. Gateways are assigned the next MAC address in sequence. A virtual forwarder that is assigned a virtual MAC address by the AVG is known as a primary virtual forwarder. Other members of the GLBP group learn the virtual MAC addresses from hello messages. A virtual forwarder that has learned the virtual MAC address is referred to as a secondary virtual forwarder.

GLBP operates virtual gateway redundancy in the same way as HSRP. One gateway is elected as the AVG, another gateway is elected as the standby virtual gateway, and the remaining gateways are placed in a listen state. If an AVG fails, the standby virtual gateway will assume responsibility for the virtual IP address. A new standby virtual gateway is then elected from the gateways in the listen state.

Virtual forwarder redundancy is similar to virtual gateway redundancy with an AVF. If the AVF fails, one of the secondary virtual forwarders in the listen state assumes responsibility for the virtual MAC address.

The new AVF is also a primary virtual forwarder for a different forwarder number. GLBP migrates hosts away from the old forwarder number using two timers that start as soon as the gateway changes to the active virtual forwarder state. GLBP uses the hello messages to communicate the current state of the timers.

The redirect time is the interval during which the AVG continues to redirect hosts to the old virtual forwarder MAC address. When the redirect time expires, the AVG stops redirecting hosts to the virtual forwarder, although the virtual forwarder will continue to forward packets that were sent to the old virtual forwarder MAC address.

The secondary hold time is the interval during which the virtual forwarder is valid. When the secondary hold time expires, the virtual forwarder is removed from all gateways in the GLBP group. The expired virtual forwarder number becomes eligible for reassignment by the AVG.

EtherChannel

On the Cisco Catalyst 4500 E-Series, EtherChannel bundles individual Ethernet links into a single logical link that provides bandwidth up to 800 Mbps (Fast EtherChannel full duplex), 8 Gbps (Gigabit EtherChannel), or 20 Gbps (10 Gigabit EtherChannel) between a Cisco Catalyst 4500 E-Series Switch and another switch or host.

A Cisco Catalyst 4500 E-Series Switch supports a maximum of 512 EtherChannel connections. You can form an EtherChannel connection with up to eight compatibly configured Ethernet, Fast Ethernet, or Gigabit Ethernet interfaces across modules in a Cisco Catalyst 4500 E-Series Switch. You can form a 10 Gigabit EtherChannel using interfaces on the Supervisor 6-E or the 6 Port 10 Gigabit Ethernet Module. All interfaces in each EtherChannel connection must be the same speed and must be configured as either Layer 2 or Layer 3 interfaces.

You can configure EtherChannel connections manually, or you can use the Port Aggregation Control Protocol (PAgP) or the Link Aggregation Control Protocol (LACP) to form EtherChannel connections. The EtherChannel protocols allow ports with similar characteristics to form an EtherChannel connection through dynamic negotiation with connected network devices. PAgP is a Cisco proprietary protocol, and LACP is defined in IEEE 802.3ad.

PAgP and LACP do not interoperate. Ports configured to use PAgP cannot form EtherChannel connections with ports configured to use LACP and vice versa.

Table 1 lists the user-configurable EtherChannel modes.

Table 1. EtherChannel Modes

Mode	Description
On	Mode that forces the LAN port to channel unconditionally. In the on mode, a usable EtherChannel connection exists only when a LAN port group in the on mode is connected to another LAN port group in the on mode. Because ports configured in the on mode do not negotiate, there is no negotiation traffic between the ports.
Auto	PAgP mode that places a LAN port into a passive negotiating state in which the port responds to PAgP packets it receives but does not initiate PAgP negotiation.
Desirable	PAgP mode that places a LAN port into an active negotiating state in which the port initiates negotiations with other LAN ports by sending PAgP packets.
Passive	LACP mode that places a port into a passive negotiating state in which the port responds to LACP packets it receives but does not initiate LACP negotiation.
Active	LACP mode that places a port into an active negotiating state in which the port initiates negotiations with other ports by sending LACP packets.

PAgP supports the automatic creation of EtherChannel connections by exchanging PAgP packets between LAN ports. PAgP packets are exchanged only between ports in **auto** and **desirable** modes.

The protocol learns the capabilities of LAN port groups dynamically and informs the other LAN ports. Once PAgP identifies correctly matched Ethernet links, it facilitates grouping the links into an EtherChannel connection. The EtherChannel connection is then added to the spanning tree as a single bridge port.

Both the **auto** and **desirable** modes allow PAgP to negotiate between LAN ports to determine if they can form an EtherChannel connection, based on criteria such as port speed and trunking state. Layer 2 EtherChannel connections also use VLAN numbers.

LAN ports can form an EtherChannel connection when they are in different PAgP modes if the modes are compatible. For example:

- A LAN port in **desirable** mode can form an EtherChannel connection successfully with another LAN port that is in **desirable** mode.
- A LAN port in **desirable** mode can form an EtherChannel connection with another LAN port in **auto** mode.

- A LAN port in **auto** mode cannot form an EtherChannel connection with another LAN port that is also in **auto** mode because neither port initiates negotiation.

LACP supports the automatic creation of EtherChannel connections by exchanging LACP packets between LAN ports. LACP packets are exchanged only between ports in **passive** and **active** modes.

The protocol learns the capabilities of LAN port groups dynamically and informs the other LAN ports. Once LACP identifies correctly matched Ethernet links, it facilitates grouping the links into an EtherChannel connection. The EtherChannel connection is then added to the spanning tree as a single bridge port.

Both the **passive** and **active** modes allow LACP to negotiate between LAN ports to determine if they can form an EtherChannel connection, based on criteria such as port speed and trunking state. Layer 2 EtherChannel connections also use VLAN numbers.

LAN ports can form an EtherChannel connection when they are in different LACP modes as long as the modes are compatible. For example:

- A LAN port in **active** mode can form an EtherChannel connection successfully with another LAN port that is in **active** mode.
- A LAN port in **active** mode can form an EtherChannel connection with another LAN port in **passive** mode.
- A LAN port in **passive** mode cannot form an EtherChannel connection with another LAN port that is also in **passive** mode, because neither port initiates negotiation.

LACP uses the following parameters:

- **LACP system priority:** You may configure an LACP system priority on each switch running LACP. The system priority can be configured automatically or through the CLI. LACP uses the system priority with the switch MAC address to form the system ID and also during negotiation with other systems.
- **LACP port priority:** You must configure an LACP port priority on each port configured to use LACP. The port priority can be configured automatically or through the CLI. See the ["Configuring Layer 2 EtherChannel Connections"](#) section. LACP uses the port priority with the port number to form the port identifier.
- **LACP administrative key:** LACP automatically configures an administrative key value equal to the channel group identification number on each port configured to use LACP. The administrative key defines the ability of a port to aggregate with other ports. A port's ability to aggregate with other ports is determined by these factors:
 - Port physical characteristics, such as data rate, duplex capability, and point-to-point or shared medium
 - Configuration restrictions that you establish

LACP tries to configure the maximum number of compatible ports in an EtherChannel connection up to the maximum allowed by the hardware (eight ports). If a port cannot be actively included in a channel, it is not included automatically if a channelled port fails.

EtherChannel can balance the traffic load across the links in the channel by reducing part of the binary pattern formed from the addresses or ports in the frame to a numerical value that selects one of the links in the channel. To balance the load, EtherChannel uses MAC addresses, IPv4 or IPv6 addresses, or Layer 4 port numbers.

Use the option that provides the greatest variety in your configuration. For example, if the traffic on a channel is going only to a single MAC address, using the destination MAC address always chooses the same link in the channel; using source addresses or IP addresses might result in better load balancing.

Power Supply Redundancy

The Cisco Catalyst 4500 E-Series chassis has two power supply bays that support two of the same supplies in a redundant mode or in a combined power-sharing mode. The supplies must be of the same type, although the system does support upgrading to a larger supply or to an inline power capable supply from a non-inline power capable supply without powering down the switch; a warning message will display during the upgrade to a larger capacity power supply. The sole exception to this is AC and DC power supplies cannot be mixed in the same chassis, even for upgrades.

There are currently five internal AC power supply options available and one external AC power shelf option for the Cisco Catalyst 4500 E-Series.

- **1000WAC:** A data only 110-220VAC power supply that can provide up to 1000W of 12VDC for line cards and supervisor engines.
- **1400WAC:** A data only 110-220VAC power supply that can provide up to 1400W of 12VDC for line cards and supervisor engines.
- **1300WAC:** A combined data and inline 110-220VAC power supply that can provide up to 1000W of 12VDC for line cards and supervisor engines, and up to 800W at -48VDC for inline power devices. The maximum combined data power and inline-powered devices cannot exceed 1300W.
- **2800WAC:** A combined data and inline 220VAC power supply that can provide 1360W of 12VDC for line cards and supervisor engines; 40W are reserved for the 3.3VDC components, such as clock oscillators. 1,400W are reserved for the -48VDC inline-powered devices.
- **4200WAC:** A combined data and inline dual 110 or 220VAC input power supply that can provide up to 1360W of power for 12VDC for line cards and supervisor engines, and 3700W for -48VDC for Power over Ethernet (PoE). This power supply is unique in the sense that it houses two individual power supplies (subunits) each providing 2100W of power. Each power supply requires individual 110 or 220VAC inputs where each input has a dedicated circuit.
- **The Cisco Catalyst 4500 external AC power shelf** is a rack-mounted, external power source that provides native DC power to the Cisco Catalyst 4500 DC power supply when operating in an AC power environment.

The Cisco Catalyst 4500 E-Series can also operate in a DC power environment. There are two options:

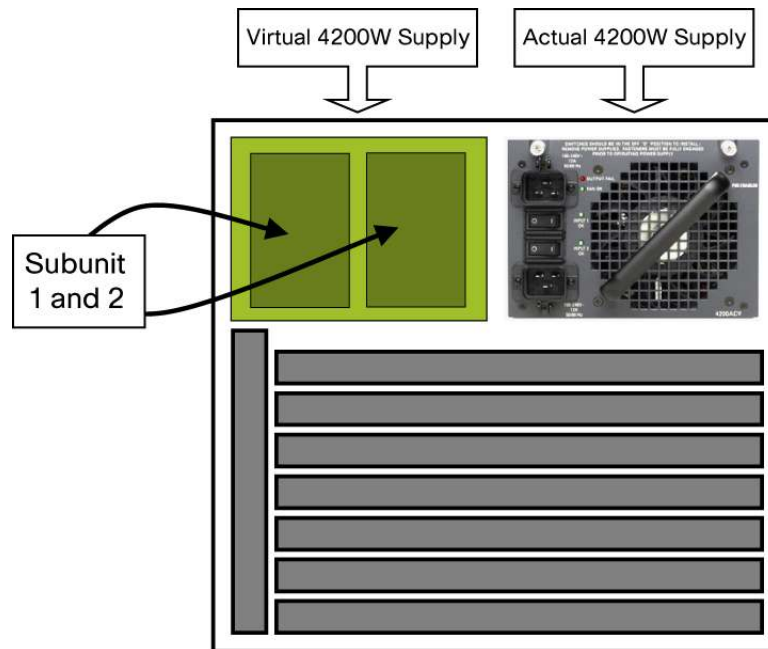
- **1400W DC-P:** This power supply is used for high-power PoE deployments. It delivers up to 1360W for the 12VDC components such as line cards with 40W reserved for the 3.3VDC components such as clock oscillators. There is also an integrated -48VDC pass-through function that enables inline power delivery through inline power capable line cards. Both current Cisco PoE implementation line cards and IEEE 802.3af compliant line cards are supported by the DC power supply. Each DC power supply is rated to provide a maximum of 7500W. The DC input power can be from any DC source, such as a battery plant, or from an AC connected source that delivers DC output power, such as the Cisco Catalyst 4500 external AC power shelf.
- **1400WDC:** This power supply is optimized for data-only deployments in the service provider central office. This triple AC input power supply allows central office technicians to customize the output power to meet their application needs.

Power Supply Configuration Modes

The Cisco Catalyst 4500 E-Series provides three different levels of power redundancy using three configuration modes:

- **Redundant mode:** In redundant mode, the second supply is online and provides half the power the system is using. The supervisor engine allocates the total available power to be less than or equal to the power available from a single supply. This provides fault coverage for the complete loss of a power supply. It is recommended that the power supplies be plugged into different power rails and preferable that they each be protected by an appropriately sized uninterruptible power supply. Redundant mode is the default mode used by the system.
- **Combined mode:** In combined mode, the supervisor engine manages the combined power budget of both supplies to provide more power than a single supply. This mode is only required for powering PoE devices that need more than the 800W or 3700W (dual 220VAC input) of PoE provided by the current 1300W and 4200W power supplies, respectively.
- **N+1 combined mode:** The N+1 combined mode (Figure 3) is an extension of the traditional combined mode and applicable only with the 4200W power supply. In this nonconfigurable mode, there are three subunits providing power to the system as a result of one of the subunits failing. This is possible because internally a single 4200W power supply consists of two subunits and in this case, one of the subunits in the primary or secondary 4200W power supply has failed. In this mode, the absolute total amount of power available for PoE and data can be up to 5500W when using 220VAC inputs and 2728W when using 110VAC inputs. This unique mode can be used as the guaranteed minimum amount of power available in combined mode by the network administrator in the event of a single subunit failure.

Figure 3. N+1 Power Mode



As with the redundant configuration, it is recommended that the power supplies be connected to an uninterruptible power supply for added resiliency. This is especially important for PoE devices because in combined mode they can use more power than a single supply can provide, and the loss of a single supply may cause some PoE devices to be disabled.

Power in combined mode does not sum linearly due to limitations in the output of each power supply. Note that the total power available in a redundant configuration is equivalent to the power of a single supply. When running in redundant mode, if a power supply fails, the remaining power supply will be able to support the Powered Devices (PDs). However, if traditional combined mode is used and a single power supply or subunit fails, resulting in N+1 combined mode if using the 4200W power supplies, the following cases can happen:

- The number of PDs actually online can be handled by a single power supply. In this case, all PDs stay up.
- The PDs are actually consuming less power at the moment of failure than the remaining power supply is providing. However, the switch has allocated more inline power than is now available. This can happen because the software allocates the maximum power a PD needs, but normally the PD is consuming less than that maximum. PDs will be shut down until the power allocated to the PDs by the software is less than the remaining inline power. "Auto" mode ports are shut down before "static" ports. If shutting down "auto" ports is insufficient, "static" ports are shut down as needed. The order of shutdown is from the bottom of the chassis to the top from and from right to left, that is, starting at interface 7/48 and ending at interface 1/1 assuming a fully populated Cisco Catalyst 4507R-E system.

- The PDs are actually consuming more power at the moment of failure than the remaining power supply is providing. *In this case, the hardware shuts down the power supply, and manual intervention is required to restart the switch.* This can also occur when the PoE power is oversubscribed in combined mode with two power supplies. The switch goes through a power evaluation cycle, with the supervisor engine and line cards being brought online first. The “static” PoE ports are brought up next, and then with the remaining power, as many “auto” mode ports as possible are brought up. This scenario will happen when the PoE requirements are manually configured and underestimated. It can also occur when rogue IEEE PDs are not providing accurate information to the Power Sourcing Equipment (PSE) as to their power requirements.

Conclusion

The Cisco Catalyst 4500 E-Series provides several features that contribute to high availability. In-Service Software Upgrade (ISSU) allows Cisco Catalyst 4500 E-Series Switches to upgrade or downgrade software with less than 200 milliseconds loss of traffic. ISSU depends on Stateful Switchover (SSO), which maintains Layer 2 adjacencies during the switchover between an active and backup supervisor engine. Nonstop Forwarding (NSF) similarly maintains Layer 3 adjacencies between the Cisco Catalyst 4500 E-Series Switch and neighboring routers during a stateful switchover.

The Supervisor Engine 6-E employs dual 10 Gigabit Ethernet uplinks. In systems with redundant supervisor engines, the uplinks on both supervisor engines can forward traffic concurrently.

The Gateway Load Balancing Protocol allows multiple Cisco Catalyst 4500 E-Series Switches to share default gateway duties with each other. Should one of the switches fail, the remaining switches transparently take over.

EtherChannel bundles up to eight links between two switches into a single logical link where packets are load-shared between the individual links.

Finally, the Cisco Catalyst 4500 E-Series provides for dual redundant power supplies to avoid power related failures.



Americas 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: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
168 Robinson Road
#28-01 Capital Tower
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Europe Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: +31 0 800 020 0791
Fax: +31 0 20 357 1100

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.

©2007 Cisco Systems, Inc. All rights reserved. CCVP, the Cisco logo, and the Cisco Square Bridge logo are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Registrar, Aironet, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, LightStream, Linksys, MeetingPlace, MGX, Networking Academy, Network Registrar, PIX, ProConnect, ScriptShare, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website 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. (0708R)