Configuring Cache Services Using the Web Cache Communication Protocol

Finding Feature Information

Your software release may not support all the features documented in this module. For the latest caveats and feature information, see Bug Search Tool and the release notes for your platform and software release. To find information about the features documented in this module, and to see a list of the releases in which each feature is supported, see the feature information table at the end of this module.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to https://cfnng.cisco.com/. An account on Cisco.com is not required.

Prerequisites for WCCP

Before configuring WCCP on your switch, make sure you adhere to the following configuration prerequisites:

- The application engines and switches in the same service group must be in the same subnetwork directly connected to the switch that has WCCP enabled.
- Configure the switch interfaces that are connected to the clients, the application engines, and the server as Layer 3 interfaces (routed ports and switch virtual interfaces [SVIs]). For WCCP packet redirection to work, the servers, application engines, and clients must be on different subnets.
- Use only nonreserved multicast addresses when configuring a single multicast address for each application engine.
- WCCP entries and PBR entries use the same TCAM region. WCCP is supported only on the templates that support PBR: access, routing, and dual IPv4/v6 routing.
- When TCAM entries are not available to add WCCP entries, packets are not redirected and are forwarded by using the standard routing tables.

- The number of available policy-based routing (PBR) labels are reduced as more interfaces are enabled for WCCP ingress redirection. For every interface that supports service groups, one label is consumed. The WCCP labels are taken from the PBR labels. You need to monitor and manage the labels that are available between PBR and WCCP. When labels are not available, the switch cannot add service groups. However, if another interface has the same sequence of service groups, a new label is not needed, and the group can be added to the interface.
- The routing maximum transmission unit (MTU) size configured on the stack member switches should be larger than the client MTU size. The MAC-layer MTU size configured on ports connected to application engines should consider the GRE tunnel header bytes.

Restrictions for WCCP

Unsupported WCCP Features

The following WCCP features are not supported in this software release:

- Packet redirection on an outbound interface that is configured by using the **ip wccp redirect out** interface configuration command.
- The GRE forwarding method for packet redirection.
- GRE redirect and return.
- On the Cisco Catalyst 3650-CX switches, to avoid packet loss you must use the flow control interface configuration command on the 1 gigabyte port connected to the Customer Edge (CE).
- WCCP over GRE
- The hash assignment method for load balancing.
- SNMP support for WCCP.
- Hash assignments in hardware. You can load balance using mask assignments only.
- Redirection for fragmented packets. This is a security feature.
- WCCP with multicast.

General Restrictions

- Maximum number of service groups: eight ingress and eight egress.
- You cannot configure WCCP and VPN routing/forwarding (VRF) on the same switch interface.
- You cannot configure WCCP and PBR on the same switch interface.
- You cannot configure WCCP and a private VLAN (PVLAN) on the same switch interface.
- The **ip wccp redirect exclude in** command allows you to exclude ingress packets from egress WCCP methods. It is not needed on the interface to CE.
- When no cache engine is available, matching packets are dropped. This is closed group support. There is no VRF-aware WCCP support and no IPv6 WCCP.

- When the device is configured with the **ip wccp check services all** command, if the redirect ACL fails to match on packet, it will be checked against the next priority service group.

Information About WCCP

WCCP Overview

Note To use this feature, the device must be running the IP Services feature set.

WCCP is supported only on Cisco Catalyst 3560-CX switches.

WCCP is a Cisco-developed content-routing technology that you can use to integrate wide-area application engines (referred to as application engines) into your network infrastructure. The application engines transparently store frequently accessed content and then fulfill successive requests for the same content, eliminating repetitive transmissions of identical content from servers. Application engines accelerate content delivery and ensure maximum scalability and availability of content. In a service-provider network, you can deploy the WCCP and application engine solution at the points of presence (POPs). In an enterprise network, you can deploy the WCCP and application engine solution at a regional site or small branch office.

The WCCP and Cisco cache engines (or other application engines running WCCP) localize traffic patterns in the network, enabling content requests to be fulfilled locally.

WCCP enables supported Cisco routers and devices to transparently redirect content requests. With transparent redirection, users do not have to configure their browsers to use a web proxy. Instead, they can use the target URL to request content, and their requests are automatically redirected to an application engine. The word transparent means that the end user does not know that a requested file (such as a web page) came from the application engine instead of from the originally specified server.

When an application engine receives a request, it attempts to service it from its own local cache. If the requested information is not present, the application engine sends a separate request to the end server to retrieve the requested information. After receiving the requested information, the application engine forwards it to the requesting client and also caches it to fulfill future requests.

With WCCP, the application-engine cluster (a series of application engines) can service multiple routers or devices.

WCCP Message Exchange

The following sequence of events describes the WCCP message exchange:

1. The application engines send their IP addresses to the WCCP-enabled device by using WCCP, signaling their presence through a Here I am message. The device and application engines communicate to each other through a control channel based on UDP port 2048.
2. The WCCP-enabled device uses the application engine IP information to create a cluster view (a list of application engines in the cluster). This view is sent through an I see you message to each application engine in the cluster, essentially making all the application engines aware of each other. A stable view is established after the membership of the cluster remains the same for a certain amount of time.

3. When a stable view is established, the application engine in the cluster with the lowest IP address is elected as the designated application engine.

WCCP Negotiation

In the exchange of WCCP protocol messages, the designated application engine and the WCCP-enabled device negotiate these items:

- Forwarding method (the method by which the device forwards packets to the application engine). The device rewrites the Layer 2 header by replacing the packet destination MAC address with the target application engine MAC address. It then forwards the packet to the application engine. This forwarding method requires the target application engine to be directly connected to the device at Layer 2.
- Assignment method (the method by which packets are distributed among the application engines in the cluster). The device uses some bits of the destination IP address, the source IP address, the destination Layer 4 port, and the source Layer 4 port to determine which application engine receives the redirected packets.
- Packet-return method (the method by which packets are returned from the application engine to the device for normal forwarding). These are the typical reasons why an application engine rejects packets and starts the packet-return feature:
  - The application engine is overloaded and has no room to service the packets.
  - The application engine receives an error message (such as a protocol or authentication error) from the server and uses the dynamic client bypass feature. The bypass enables clients to bypass the application engines and to connect directly to the server.

The application engine returns a packet to the WCCP-enabled device to forward to the server as if the application engine is not present. The application engine does not intercept the reconnection attempt. In this way, the application engine effectively cancels the redirection of a packet to the application engine and creates a bypass flow. If the return method is Layer 2 rewrite, the packets are forwarded in hardware to the target server. When the server responds with the information, the device uses normal Layer 3 forwarding to return the information to the requesting client.

MD5 Security

WCCP provides an optional security component in each protocol message to enable the device to use MD5 authentication on messages between the device and the application engine. Messages that do not authenticate by MD5 (when authentication of the device is enabled) are discarded by the device. The password string is combined with the MD5 value to create security for the connection between the device and the application engine. You must configure the same password on each application engine.

Packet Redirection and Service Groups

You can configure WCCP to classify traffic for redirection, such as FTP, proxy-web-cache handling, and audio and video applications. This classification, known as a service group, is based on the protocol type (TCP or UDP) and the Layer 4 source destination port numbers. The service groups are identified either by well-known names such as web-cache, which means TCP port 80, or a service number, 0 to 99. Service groups are configured to map to a protocol and Layer 4 port numbers and are established and maintained independently.

WCCP allows dynamic service groups, where the classification criteria are provided dynamically by a participating application engine.

You can configure up to 8 service groups on a device or device stack and up to 32 cache engines per service group. WCCP maintains the priority of the service group in the group definition. WCCP uses the priority to configure the service groups in the device hardware. For example, if service group 1 has a priority of 100 and looks for destination port 80, and service group 2 has a priority of 50 and looks for source port 80, the incoming packet with source and destination port 80 is forwarded by using service group 1 because it has the higher priority.

WCCP supports a cluster of application engines for every service group. Redirected traffic can be sent to any one of the application engines. The device supports the mask assignment method of load balancing the traffic among the application engines in the cluster for a service group.

After WCCP is configured on the device, the device forwards all service group packets received from clients to the application engines. However, the following packets are not redirected:

- Packets originating from the application engine and targeted to the server.
- Packets originating from the application engine and targeted to the client.
- Packets returned or rejected by the application engine. These packets are sent to the server.

You can configure a single multicast address per service group for sending and receiving protocol messages. When there is a single multicast address, the application engine sends a notification to one address, which provides coverage for all routers in the service group, for example, 225.0.0.0. If you add and remove routers dynamically, using a single multicast address provides easier configuration because you do not need to specifically enter the addresses of all devices in the WCCP network.

You can use a router group list to validate the protocol packets received from the application engine. Packets matching the address in the group list are processed, packets not matching the group list address are dropped.

To disable caching for specific clients, servers, or client/server pairs, you can use a WCCP redirect access control list (ACL). Packets that do not match the redirect ACL bypass the cache and are forwarded normally.

Before WCCP packets are redirected, the device examines ACLs associated with all inbound features configured on the interface and permits or denies packet forwarding based on how the packet matches the entries in the ACL.

Note Both permit and deny ACL entries are supported in WCCP redirect lists.

When packets are redirected, the output ACLs associated with the redirected interface are applied to the packets. Any ACLs associated with the original port are not applied unless you specifically configure the required output ACLs on the redirected interfaces.

How to Configure WCCP

Default WCCP Configuration

Feature	Default Setting
WCCP enable state	WCCP services are disabled.
Protocol version	WCCPv2.
Redirecting traffic received on an interface	Disabled.

Enabling the Cache Service

For WCCP packet redirection to operate, you must configure the device interface connected to the client to redirect inbound packets.

This procedure shows how to configure these features on routed ports. To configure these features on SVIs, see the configuration examples that follow the procedure.

Follow these steps to enable the cache service, to set a multicast group address or group list, to configure routed interfaces, to redirect inbound packets received from a client to the application engine, enable an interface to listen for a multicast address, and to set a password. This procedure is required.

Before you begin

Configure the SDM template, and reboot the device.

SUMMARY STEPS

1. **enable**
2. **configure terminal**
3. **ip wccp** {**web-cache** | *service-number*} [**group-address** *groupaddress*] [**group-list** *access-list*] [**redirect-list** *access-list*] [**password** *encryption-number password*]
4. **interface** *interface-id*
5. **no switchport**
6. **ip address** *ip-address subnet-mask*
7. **no shutdown**
8. **exit**
9. **interface** *interface-id*
10. **no switchport**
11. **ip address** *ip-address subnet-mask*
12. **no shutdown**
13. **ip wccp** {**web-cache** | *service-number*} **redirect in**
14. **ip wccp** {**web-cache** | *service-number*} **group-listen**
15. **exit**
16. **end**
17. **show running-config**

18. **copy running-config startup-config**

DETAILED STEPS

	Command or Action	Purpose
Step 1	**enable** **Example:** `Switch> enable`	Enables privileged EXEC mode. • Enter your password if prompted.
Step 2	**configure terminal** **Example:** `Switch# configure terminal`	Enters global configuration mode.
Step 3	**ip wccp** {**web-cache** | *service-number*} [**group-address** *groupaddress*] [**group-list** *access-list*] [**redirect-list** *access-list*] [**password** *encryption-number password*] **Example:** `Switch(config)# ip wccp web-cache`	Enables the cache service, and specifies the service number that corresponds to a dynamic service that is defined by the application engine. By default, this feature is disabled. (Optional) For **group-address** *groupaddress*, specifies the multicast group address used by the devices and the application engines to participate in the service group. (Optional) For **group-list** *access-list*, if a multicast group address is not used, specify a list of valid IP addresses that correspond to the application engines that are participating in the service group. (Optional) For **redirect-list** *access-list*, specify the redirect service for specific hosts or specific packets from hosts. (Optional) For **password** *encryption-number password*, specify an encryption number. The range is 0 to 7. Use 0 for not encrypted, and use 7 for proprietary. Specify a password name up to seven characters in length. The device combines the password with the MD5 authentication value to create security for the connection between the device and the application engine. By default, no password is configured, and no authentication is performed. You must configure the same password on each application engine. When authentication is enabled, the device discards messages that are not authenticated.
Step 4	**interface** *interface-id* **Example:** `Switch(config)# interface gigabitethernet1/0/1`	Specifies the interface connected to the application engine or the server, and enters interface configuration mode.

	Command or Action	Purpose
Step 5	**no switchport** **Example:** `Switch(config-if)# no switchport`	Enters Layer 3 mode.
Step 6	**ip address** *ip-address subnet-mask* **Example:** `Switch(config-if)# ip address 172.20.10.30 255.255.255.0`	Configures the IP address and subnet mask.
Step 7	**no shutdown** **Example:** `Switch(config-if)# no shutdown`	Enables the interface.
Step 8	**exit** **Example:** `Switch(config-if)# exit`	Returns to global configuration mode. Repeat Steps 4 through 8 for each application engine and server.
Step 9	**interface** *interface-id* **Example:** `Switch(config)# interface gigabitethernet1/0/2`	Specifies the interface connected to the client, and enters interface configuration mode.
Step 10	**no switchport** **Example:** `Switch(config-if)# no switchport`	Enters Layer 3 mode.
Step 11	**ip address** *ip-address subnet-mask* **Example:** `Switch(config-if)# ip address 175.20.20.10 255.255.255.0`	Configures the IP address and subnet mask.
Step 12	**no shutdown** **Example:** `Switch(config-if)# no shutdown`	Enables the interface.
Step 13	**ip wccp** {**web-cache** | *service-number*} **redirect in** **Example:** `Switch(config-if)# ip wccp web-cache redirect in`	Redirects packets received from the client to the application engine. Enable this on the interface connected to the client.

	Command or Action	Purpose
Step 14	**ip wccp** {**web-cache** | *service-number*} **group-listen** **Example:** `Switch(config-if)# ip wccp web-cache group-listen`	(Optional) When using a multicast group address, the **group-listen** keyword enables the interface to listen for the multicast address. Enable this on the interface connected to the application engine.
Step 15	**exit** **Example:** `Switch(config-if)# exit`	Returns to global configuration mode. Repeat Steps 9 through 15 for each client.
Step 16	**end** **Example:** `Switch(config)# end`	Returns to privileged EXEC mode.
Step 17	**show running-config** **Example:** `Switch# show running-config`	Verifies your entries.
Step 18	**copy running-config startup-config** **Example:** `Switch# copy running-config startup-config`	(Optional) Saves your entries in the configuration file.

Configuration Examples

This example shows how to configure routed interfaces and to enable the cache service with a multicast group address and a redirect access list. Gigabit Ethernet port 1 is connected to the application engine, is configured as a routed port with an IP address of 172.20.10.30, and is reenabled. Gigabit Ethernet port 2 is connected through the Internet to the server, is configured as a routed port with an IP address of 175.20.20.10, and is reenabled. Gigabit Ethernet ports 3 to 6 are connected to the clients and are configured as routed ports with IP addresses 175.20.30.20, 175.20.40.30, 175.20.50.40, and 175.20.60.50. The device listens for multicast traffic and redirects packets received from the client interfaces to the application engine.

```
Switch# configure terminal
Switch(config)# ip wccp web-cache group-address 224.1.1.100 redirect list 12
Switch(config)# access-list 12 permit host 10.1.1.1
Switch(config)# interface gigabitethernet1/0/1
Switch(config-if)# no switchport
Switch(config-if)# ip address 172.20.10.30 255.255.255.0
Switch(config-if)# no shutdown
Switch(config-if)# ip wccp web-cache group-listen
Switch(config-if)# exit
Switch(config)# interface gigabitethernet1/0/2
```

```
Switch(config-if)# no switchport
Switch(config-if)# ip address 175.20.20.10 255.255.255.0
Switch(config-if)# no shutdown
Switch(config-if)# exit
Switch(config)# interface gigabitethernet1/0/3
Switch(config-if)# no switchport
Switch(config-if)# ip address 175.20.30.20 255.255.255.0
Switch(config-if)# no shutdown
Switch(config-if)# ip wccp web-cache redirect in
Switch(config-if)# exit
Switch(config)# interface gigabitethernet1/0/4
Switch(config-if)# no switchport
Switch(config-if)# ip address 175.20.40.30 255.255.255.0
Switch(config-if)# no shutdown
Switch(config-if)# ip wccp web-cache redirect in
Switch(config-if)# exit
Switch(config)# interface gigabitethernet1/0/5
Switch(config-if)# no switchport
Switch(config-if)# ip address 175.20.50.40 255.255.255.0
Switch(config-if)# no shutdown
Switch(config-if)# ip wccp web-cache redirect in
Switch(config-if)# exit
Switch(config)# interface gigabitethernet1/0/6
Switch(config-if)# no switchport
Switch(config-if)# ip address 175.20.60.50 255.255.255.0
Switch(config-if)# no shutdown
Switch(config-if)# ip wccp web-cache redirect in
Switch(config-if)# exit
```

This example shows how to configure SVIs and how to enable the cache service with a multicast group list. VLAN 299 is created and configured with an IP address of 175.20.20.10. Gigabit Ethernet port 1 is connected through the Internet to the server and is configured as an access port in VLAN 299. VLAN 300 is created and configured with an IP address of 172.20.10.30. Gigabit Ethernet port 2 is connected to the application engine and is configured as an access port in VLAN 300. VLAN 301 is created and configured with an IP address of 175.20.30.50. Fast Ethernet ports 3 to 6, which are connected to the clients, are configured as access ports in VLAN 301. The device redirects packets received from the client interfaces to the application engine.

Note Both permit and deny ACL entries are supported in WCCP redirect lists.

```
Switch# configure terminal
Switch(config)# ip wccp web-cache group-list 15
Switch(config)# access-list 15 permit host 171.69.198.102
Switch(config)# access-list 15 permit host 171.69.198.104
Switch(config)# access-list 15 permit host 171.69.198.106
Switch(config)# vlan 299
Switch(config-vlan)# exit
Switch(config)# interface vlan 299
Switch(config-if)# ip address 175.20.20.10 255.255.255.0
Switch(config-if)# exit
Switch(config)# interface gigabitethernet1/0/1
Switch(config-if)# switchport mode access
Switch(config-if)# switchport access vlan 299
Switch(config)# vlan 300
Switch(config-vlan)# exit
Switch(config)# interface vlan 300
Switch(config-if)# ip address 171.69.198.100 255.255.255.0
Switch(config-if)# exit
```

```
Switch(config)# interface gigabitethernet1/0/2
Switch(config-if)# switchport mode access
Switch(config-if)# switchport access vlan 300
Switch(config-if)# exit
Switch(config)# vlan 301
Switch(config-vlan)# exit
Switch(config)# interface vlan 301
Switch(config-if)# ip address 175.20.30.20 255.255.255.0
Switch(config-if)# ip wccp web-cache redirect in
Switch(config-if)# exit
Switch(config)# interface range gigabitethernet1/0/3 - 6
Switch(config-if-range)# switchport mode access
Switch(config-if-range)# switchport access vlan 301
Switch(config-if-range)# exit
```

What to do next

To disable the cache service, use the **no ip wccp web-cache** global configuration command. To disable inbound packet redirection, use the **no ip wccp web-cache redirect in** interface configuration command. After completing this procedure, configure the application engines in the network.