Network Working Group
Internet-Draft
Intended status: Informational
Expires: February 3, 2013

D. McLaggan
Cisco Systems

August 2, 2012

Web Cache Communication Protocol V2, Revision 1
draft-mclaggan-wccp-v2rev1-00

Abstract

This document describes version 2 of the Web Cache Communication Protocol (WCCP). The WCCP V2 protocol specifies interactions between one or more routers and one or more web-caches. The interaction may take place within an IPv4 or IPv6 network. The purpose of the interaction is to establish and maintain the transparent redirection of selected types of traffic flowing through a group of routers (or similar devices). The selected traffic is redirected to a group of web-caches (or other traffic optimisation devices) with the aim of optimising resource usage and lowering response times.

The protocol does not specify any interaction between the web-caches within a group or between a web-cache and a web-server.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79. This document may not be modified, and derivative works of it may not be created, except to format it for publication as an RFC or to translate it into languages other than English.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 3, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Table of Contents

1. Introduction

1.1. Protocol Overview

WCCP V2 defines mechanisms to allow one or more routers enabled for transparent redirection to discover, verify, and advertise connectivity to one or more web-caches.

Having established connectivity the routers and web-caches form Service Groups to handle the redirection of traffic whose characteristics are part of the Service Group definition.

The protocol provides the means to negotiate the specific method used for load distribution among web-caches and also the method used to transport traffic between a router and a web-cache.

A single web-cache within a Service Group is elected as the designated web-cache. It is the responsibility of the designated web-cache to provide routers with the data which determines how redirected traffic is distributed between the web-caches in the Service Group.

Although its original purpose was for use with web-caches, the WCCP V2 protocol is suitable for use with many types of network devices that need to transparently intercept IP traffic. For the sake of simplicity and to maintain consistency with the protocol name, the device wishing to receive redirected IP traffic will be generically referred to as the "web-cache" in this document.

Similarly, the device through which the IP traffic to be redirected is flowing will generically be referred to in this document as the "router", even though the protocol is suitable for use with several types of network devices through which IP traffic may flow.

This document specifies WCCP V2 for use with multiple address families, specifically including both IPv4 and IPv6. References here to "IP" apply equally to both IPv4 and IPv6 and are used when the discussion is not specific to a particular address family.

1.2. Contributing Authors

This document is derived from the work of the following authors who wrote the original description of WCCP Version 2 in July 2000:

* Martin Cieslak (Cisco Systems)
* David Forster (Cisco Systems)
* Gurumukh Tiwana (Cisco Systems)
* Rob Wilson (Cisco Systems)

The protocol described in the current document is a fully backwards-compatible extension of the originally described protocol, with extensions added to support the IPv6 address family, configurable message interval timing, more compact message formats and some additional minor enhancements.

The work of the original authors represents a very significant proportion of the current document and authorship of the majority of the protocol remains with the four authors listed above.

2. Definitions

Assignment Method

The method by which redirected packets are distributed between web-caches. Hash assignment or mask assignment can be used.

Designated Web-Cache

The web-cache in a web-cache farm responsible for dictating to the router or routers how redirected traffic should be distributed between the members of the farm.

Forwarding Method

The method by which redirected packets are transported from router to web-cache.

Packet Return Method

The method by which packets redirected to a web-cache are returned to a router for normal forwarding.

Redirection Hash Table

A 256-bucket hash table maintained by the router or routers when using hash assignment. This table maps the hash index derived from a packet to be redirected to the IP address of a destination web-cache.

Reserved

Parts of a message defined as reserved must be set to zero by the sender and must be ignored by the receiver.

Router

This term is used generically throughout this document to refer to a network device that may use the protocol to establish redirection of traffic flowing through it.

Service Group

A group of one or more routers plus one or more web-caches working together in the redirection of traffic whose characteristics are part of the Service Group definition.

Transparent Redirection

Transparent redirection is a technique used to deploy traffic optimisation without the need for reconfiguration of clients or servers. It involves the interception and redirection of traffic to one or more intervening devices by a router or switch transparently to the end points of the traffic flow.

Usable Web-Cache

From the viewpoint of a router a web-cache is considered a usable member of a Service Group when it has sent that web-cache a WCCP2_I_SEE_YOU message and has received in response a WCCP2_HERE_I_AM message with a valid "Receive ID" and compatible capabilities.

Web-cache

This term is used generically throughout this document to refer to a network device that will receive redirected traffic. The term comes from the protocol's original purpose of redirecting HTTP requests to a caching device.

Web-Cache Farm

One or more web-caches associated with a router or routers.

2.1. Time Interval Definitions

TRANSMIT_T

The time interval at which a web-cache must send successive WCCP2_HERE_I_AM messages. The default interval is 10 seconds.

TIMEOUT_BASE_T

A time interval used as the basis for calculating timeout values. The default interval is 10 seconds. The value is calculated using this formula: TIMEOUT_BASE_T = (TIMEOUT_SCALE * TRANSMIT_T).

RA_TIMER_BASE_T

A time interval used as the basis for calculating timeout values. The default interval is 10 seconds. The value is calculated using this formula: RA_TIMER_BASE_T = (RA_TIMER_SCALE * TRANSMIT_T).

TIMEOUT_SCALE

A multiplier used to calculate the value of TIMEOUT_BASE_T from the value of TRANSMIT_T. The default value of the multiplier is 1.

RA_TIMER_SCALE

A multiplier used to calculate the value of RA_TIMER_BASE_T from the value of TRANSMIT_T. The default value of the multiplier is 1.

3. Protocol Description

3.1. Joining a Service Group

A web-cache joins and maintains its membership of a Service Group by transmitting a WCCP2_HERE_I_AM message to each router in the Group at time intervals of TRANSMIT_T. This may be by unicast to each router or multicast to the configured Service Group multicast address. The Web-Cache Info Component in the WCCP2_HERE_I_AM message identifies the web-cache by IP address. The Service Info Component of the WCCP2_HERE_I_AM message identifies and describes the Service Group in which the web-cache wishes to participate.

A router responds to a WCCP2_HERE_I_AM message with a WCCP2_I_SEE_YOU message. If the WCCP2_HERE_I_AM message was unicast then the router will respond immediately with a unicast WCCP2_I_SEE_YOU message. If the WCCP2_HERE_I_AM message was multicast the router will respond later via the scheduled multicast WCCP2_I_SEE_YOU message for the Service Group.

A router responds to multicast web-cache members of a Service Group using a multicast WCCP2_I_SEE_YOU message transmitted at time intervals of 0.9 * TRANSMIT_T with a 10% jitter.

The Router Identity Component in a WCCP2_I_SEE_YOU message includes a list of the web-caches to which the packet is addressed. A web-cache not in the list should discard the WCCP2_I_SEE_YOU message.

The default value for the TRANSMIT_T interval is 10 seconds. A change in this value is only permissible if a new value is negotiated between a router and a web-cache via the WCCP2_TRANSMIT_T capability. A router or web-cache must use the value for TRANSMIT_T specified in the router's WCCP2_I_SEE_YOU message, or use the default value if a specific value has not yet been given in a WCCP2_I_SEE_YOU message. If a specific timer value has been negotiated between a web-cache and a router, the web-cache must only send HERE_I_AM messages at the negotiated interval. Support for the default 10 seconds TRANSMIT_T interval is mandatory. Support for other values of TRANSMIT_T is optional. The range of supported values may be chosen by the implementation.

Before negotiation of a non-default TRANSMIT_T interval has taken place, a web-cache may choose to send WCCP2_HERE_I_AM messages at a shorter interval than the default TRANSMIT_T interval, provided that all of the following conditions are met:

(1) all other timing calculations remain based on the default time interval of 10 seconds,

(2) the web-cache has received a WCCP2_I_SEE_YOU message containing a WCCP2_TRANSMIT_T capability describing the range of values supported by the router,

(3) the web-cache's chosen interval falls within the range supported by the router, and

(4) the negotiation of a specific WCCP2_TRANSMIT_T value has not yet completed.

3.2. Describing a Service Group

The Service Info Component of a WCCP2_HERE_I_AM message describes the Service Group in which a web-cache wishes to participate. A Service Group is identified by its Service Type and Service ID. There are two types of Service Group:

* Well Known Services
* Dynamic Services

Well Known Services are known by both routers and web-caches and do not require a description other than the Service ID. The characteristics of the traffic associated with a Well Known Service are fixed and implicitly known to both router and web-cache.

The traffic characteristics associated with a Dynamic Service are not known in advance to the router and must be described by each web-cache. A router is configured to participate in a particular Dynamic Service Group, identified by its Service ID, initially without any knowledge of the characteristics of the traffic associated with the Service Group. The traffic description is communicated to the router in the WCCP2_HERE_I_AM message of the first web-cache to join the Service Group. A web-cache describes a Dynamic Service using the Protocol, Service Flags and Port fields of the Service Info Component. Once a Dynamic Service has been defined, a router will discard any subsequent WCCP2_HERE_I_AM message which contains a conflicting description. The service definition is reset by the router when all web-caches have left the Service Group. A router will also discard any WCCP2_HERE_I_AM message which describes a Service Group for which the router has not been configured.

3.3. Establishing Two-Way Connectivity

WCCP V2 uses a "Receive ID" to verify two-way connectivity between a router and a web-cache. The Router Identity Info Component of a WCCP2_I_SEE_YOU message contains a "Receive ID" within the Router Identity Element. This value is maintained separately for each Service Group and it is incremented each time the router sends a WCCP2_I_SEE_YOU message for the Service Group. The router records the "Receive ID" value it sends to each web-cache.

The "Receive ID" sent by a router is usually reflected back by a web-cache using a Router Identity Element within the Web-Cache View Info Component of a WCCP2_HERE_I_AM message. However, when a web-cache first attempts to contact a router, no "Receive ID" will be available and the router will not be listed in the Web-Cache View Info Component.

A router checks the value given for its own "Receive ID" in each WCCP2_HERE_I_AM message received from a web-cache. The "Receive ID" is invalid if the value does not match the "Receive ID" in the most recent WCCP2_I_SEE_YOU message sent to the web-cache, or the router is not listed in Web-Cache View Info Component, or the router has not previously sent a message to the web-cache.

When the "Receive ID" is found to be invalid, the router replies with a WCCP2_I_SEE_YOU message to advertise the correct "Receive ID", but the WCCP2_HERE_I_AM message is then discarded and it is not treated as a validly received WCCP2_HERE_I_AM message. In this case most of the WCCP2_HERE_I_AM message is ignored by the router.

A router can only begin to consider a web-cache as a potentially usable member of a Service Group after it has sent that web-cache a WCCP2_I_SEE_YOU message and subsequently received a WCCP2_HERE_I_AM message from it containing the correct "Receive ID".

3.4. Negotiating the Protocol Version Number

WCCP V2 is an extensible protocol and may incorporate a number of revisions to the message format. Higher revision levels may introduce new message components, elements and formats that may not be valid at a lower revision level.

The protocol version is specified within each WCCP V2 message and consists of the major version number, which is always set to 2, combined with the minor version number, which indicates the revision level of the V2 protocol. In the context of this document, as the major version number is fixed, references to different protocol version numbers refer specifically to differences in the minor protocol version number only.

A router or web-cache may use the protocol version within a WCCP message to decide how to process or respond to an incoming message, or to indicate via an outgoing message which protocol version it supports.

A router or web-cache receiving a WCCP message should aim to process the valid components and elements of the message even if other parts of the message may not be understood or appear invalid. However, unless performing protocol version negotiation, a router or web-cache is permitted to ignore messages in which the protocol version number is not recognised.

A router or web-cache may support a single protocol version or multiple protocol versions. To support multiple versions, the router or web-cache must support negotiation of the protocol version number. The negotiation takes place per Service Group. Thus routers and web-caches participating in several Service Groups may negotiate a different protocol version for each Service Group.

A router and web-cache that communicate with each other must learn which version of the protocol is supported by the intended recipient. They should not send a message without knowing that the intended recipient can understand the message format used. The version supported by the intended recipient is determined from the protocol version set within the message most recently received from it.

The format of a message must always conform to the protocol version number set within the message header.

3.4.1. Responsibilities of a web-cache during version negotiation

When a web-cache sends the first WCCP2_HERE_I_AM message to a router, the web-cache must decide the protocol version number to use in the message without knowing which protocol versions the router is capable of supporting or understanding.

In this situation, a web-cache not wishing to negotiate the protocol version number should set the V bit to 0 within the Web-Cache Identity Element in the first WCCP2_HERE_I_AM message and set the protocol version number in the message header to the only version number that the web-cache is able to support.

Alternatively, a web-cache wishing to negotiate the protocol version should set the V bit to 1 within the Web-Cache Identity Element in the first WCCP2_HERE_I_AM message and set the protocol version number in the message header to the lowest version number that the web-cache is able to support. The lowest version number is used in this case to maximise the chance that a router will understand and respond to the message. The web-cache should only set the V bit to 1 in a WCCP2_HERE_I_AM message when it has not yet received a response from the router.

When a web-cache receives a first WCCP2_I_SEE_YOU message from a router, this provides it with information about the protocol version the router is able to support. Even if the web-cache does not support the version used by the router, the web-cache should set the V bit to 0 in subsequent WCCP2_HERE_I_AM messages and use a version number that is less than or equal to the version number the router responded with.

A web-cache need not use the V bit to negotiate the protocol version number, but using the V bit will increase the likelihood that negotiation will be successful by increasing the chance that a response will be received to the initial message.

If the V bit is not used, limited version negotiation may still take place although successful negotiation is not guaranteed as some routers may decide not to respond. In this situation the web-cache begins negotiations by setting the protocol version number within the first WCCP2_HERE_I_AM message to be the highest protocol version number supported by the web-cache. If a router replies, the response will contain either the same or a lower version number. The web-cache must then use the version number set by the router, or ignore the response from the router.

3.4.2. Responsibilities of a router during version negotiation

A router that finds the V bit set to 1 in an incoming WCCP2_HERE_I_AM message must reply by setting the protocol version number in its WCCP2_I_SEE_YOU message to the highest version it can support. In a multicast service group when a router is responding to multiple WCCP2_HERE_I_AM messages, the V bit must be set to 1 in all incoming messages before it is acted upon.

When the V bit of an incoming message is set to 0, a router must treat the protocol version number in a WCCP2_HERE_I_AM message as the maximum version the web-cache is capable of supporting. In this case a router has the option of replying using the same version number, replying using a lower version number, or not replying at all. When replying, the router responds with a version that is less than or equal to the version the web-cache used. Therefore the router may respond to the message even if it does not support the version set by the web-cache.

3.5. Negotiating Capabilities

WCCP includes a number of optional features or capabilities that an implementation may choose to support. To allow a router and web-cache to agree on which optional capabilities can be used for a particular Service Group, the capabilities are negotiated after a router's "Receive ID" has been successfully echoed back from the web-cache to the router.

For each defined capability, an implementation must support at least one option from the range of possible options defined for that particular capability. Negotiation of each capability is optional. For each capability there is a default setting which is used if negotiation of the capability does not take place. Negotiation takes place independently for each Service Group.

Currently, the following capabilities can be negotiated:

* Forwarding Method (Default: GRE encapsulation)

  The method by which packets are forwarded to a web-cache by a router.

* Assignment Method (Default: Hash assignment)

  The method by which packets are distributed between the web-caches in a Service Group.

* Packet Return Method (Default: GRE encapsulation)

  The method by which packets are returned from a web-cache to a router for normal forwarding.

* TRANSMIT_T Message Interval (Default: 10 seconds)

  The required interval between successive HERE_I_AM messages.

* TIMEOUT_SCALE and RA_TIMER_SCALE values (Default: 1 and 1)

  Two scaling factors used in message timeout calculations.

Capability negotiation requires the router to advertise the options that it currently supports for each capability of a Service Group using the optional Capabilities Info Component of the WCCP2_I_SEE_YOU message. The absence of this component implies the router supports only the default option for all capabilities. Similarly, the absence of an individual capability from within this component implies the router supports only the default option for that capability.

Negotiation with a router takes place independently for each web-cache, but the options advertised by the router may be influenced by previous negotiations with other web-caches. So, for a given Service Group, the router may permit different options to be negotiated by different web-caches, or it may force all web-caches to agree on a common option. A web-cache participating in several Service Groups may negotiate different capability options for each Service Group.

A web-cache will inspect the capabilities advertisement in the first WCCP2_I_SEE_YOU message received from a router for a particular Service Group. If the router does not advertise an option supported by the web-cache for every known capability then the web-cache will abort its attempt to join the Service Group. Otherwise the web-cache will pick one option from those advertised by the router for each capability and specify them in the optional Capabilities Info Component of its next WCCP2_HERE_I_AM message. The absence of this component in a WCCP2_HERE_I_AM message implies the web-cache is requesting the default option for all capabilities. Similarly, the absence of an individual capability from within this component implies the web-cache is requesting the default setting for that capability.

A router will inspect the capability options selected by a web-cache in a WCCP2_HERE_I_AM message, provided that the message contains a valid "Receive ID". If all of the requested options are supported, the router will accept the web-cache as usable and add it to the Service Group. Otherwise, if any of the selected options are not supported by the router, the router will not add the web-cache to the Service Group and will instead decide that the web-cache is unusable. In both cases the router will respond to the WCCP2_HERE_I_AM message, either indicating the capability options that have been successfully negotiated, or again advertising the capability options that are available.

Note that, for each Service Group, the web-cache need not include a Capabilities Info Component in a WCCP2_HERE_I_AM message until after the first WCCP2_I_SEE_YOU message from the router has been received. Following negotiation, both web-cache and router should continue to include the negotiated capabilities in every WCCP2_HERE_I_AM and WCCP2_I_SEE_YOU message. If a router or web-cache encounters an unrecognised capability at any time it should simply be ignored to allow the default setting for the capability to be selected.

3.5.1. Negotiating the Forwarding Method

A web-cache and router may negotiate the method by which packets are forwarded to the web-cache by the router.

A router will advertise the supported forwarding methods for a Service Group. The absence of such an advertisement implies the router supports the default GRE encapsulation method only.

If the router does not advertise a packet return method supported by the web-cache then the web-cache will abort its attempt to join the Service Group. Otherwise the web-cache will select a packet return method to be indicated in the next WCCP2_HERE_I_AM message. Absence of an advertisement of the forwarding method in a WCCP2_HERE_I_AM message implies the web-cache is requesting the default GRE encapsulation method.

3.5.2. Negotiating the Assignment Method

A web-cache and router may negotiate the method by which packets are distributed between the web-caches in a Service Group.

A router will advertise the supported assignment methods for a Service Group. The absence of such an advertisement implies the router supports the default Hash assignment method only.

If the router does not advertise an assignment method supported by the web-cache then the web-cache will abort its attempt to join the Service Group. Otherwise the web-cache will select an assignment method to be indicated in the next WCCP2_HERE_I_AM message. Absence of an assignment method advertisement in a WCCP2_HERE_I_AM message implies the web-cache is requesting the default Hash assignment method.

If the assignment method selected by a web-cache is supported and other capabilities have been successfully negotiated, the router will accept the web-cache as usable and add it to the Service Group. When the first web-cache joins a Service Group, the router will set the assignment method selected by the web-cache to be the only assignment method supported by the Service Group. This assignment method will remain selected until all web-caches are removed from the Service Group.

3.5.3. Negotiating the Packet Return Method

A web-cache and router may negotiate the method by which packets are returned from the web-cache to the router for normal forwarding.

A router will advertise the supported packet return methods for a Service Group. The absence of such an advertisement implies the router supports the default GRE encapsulation method only.

If the router does not advertise a packet return method supported by the web-cache then the web-cache will abort its attempt to join the Service Group. Otherwise the web-cache will select a packet return method to be indicated in the next WCCP2_HERE_I_AM message. Absence of an advertisement of the packet return method in a WCCP2_HERE_I_AM message implies the web-cache is requesting the default GRE encapsulation method.

3.5.4. Negotiating the TRANSMIT_T Message Interval Value

A web-cache and router may negotiate the TRANSMIT_T message interval value used by the Service Group.

A router will advertise the range of supported TRANSMIT_T message interval values. The range is given by specifying its upper and lower limits, or by specifying a single value.

The absence of such an advertisement implies the router supports the default TRANSMIT_T message interval of 10 seconds only. In this case the web-cache must never attempt to specify or use an alternative TRANSMIT_T message interval.

If the router does not advertise a TRANSMIT_T message interval supported by the web-cache then the web-cache will abort its attempt to join the Service Group. Otherwise the web-cache will select an interval value either within the advertised range, or matching the single advertised value. The selected value will be indicated in the next WCCP2_HERE_I_AM message. Absence of a TRANSMIT_T message interval advertisement in a WCCP2_HERE_I_AM message implies the web-cache is requesting the default TRANSMIT_T message interval of 10 seconds.

If the interval selected by a web-cache is supported and other capabilities have been successfully negotiated, the router will accept the web-cache as usable and add it to the Service Group. When the first web-cache joins a Service Group, the router will set the TRANSMIT_T message interval value selected by the web-cache to be the only value supported by the Service Group. This value will remain selected until all web-caches are removed from the Service Group.

3.5.5. Negotiating the TIMEOUT_SCALE and RA_TIMER_SCALE values

A web-cache and router may negotiate the TIMEOUT_SCALE and RA_TIMER_SCALE values used by the Service Group. Both values are negotiated together as a pair.

A router will advertise the ranges of supported TIMEOUT_SCALE values and the range of supported RA_TIMER_SCALE values for a Service Group. Each range is given by specifying its upper and lower limits, or by specifying a single value.

The absence of such an advertisement implies the router supports only the default value of 1 for both the TIMEOUT_SCALE and RA_TIMER_SCALE parameters. In this case the web-cache must never attempt to specify or use alternative TIMEOUT_SCALE and RA_TIMER_SCALE values.

If the router does not advertise TIMEOUT_SCALE and RA_TIMER_SCALE values supported by the web-cache then the web-cache will abort its attempt to join the Service Group. Otherwise the web-cache will select a TIMEOUT_SCALE value and an RA_TIMER_SCALE value, either within the advertised range, or matching the single advertised value. The selected values will be indicated in the next WCCP2_HERE_I_AM message. Absence of an advertisement of TIMEOUT_SCALE and RA_TIMER_SCALE values in a WCCP2_HERE_I_AM message implies the web-cache is requesting the default value of 1 for both the TIMEOUT_SCALE and RA_TIMER_SCALE parameters.

If the values selected by a web-cache are supported and other capabilities have been successfully negotiated, the router will accept the web-cache as usable and add it to the Service Group. When the first web-cache joins a Service Group, the router will set the TIMEOUT_SCALE and RA_TIMER_SCALE values selected by the web-cache to be the only values supported by the Service Group. These values will remain selected until all web-caches are removed from the Service Group.

3.6. Advertising Views of the Service Group

Each router advertises its view of a Service Group via the Router View Info Component in the WCCP2_I_SEE_YOU message it sends to web-caches. This component includes a list of the useable web-caches in the Service Group as seen by the router and a list of the routers in the Service Group as reported in WCCP2_HERE_I_AM messages from web-caches. A change number in the component is incremented if the Service Group membership has changed since the previous WCCP2_I_SEE_YOU message sent by the router.

Each web-cache advertises its view of the Service Group via the Web-Cache View Info Component in the WCCP2_HERE_I_AM message it sends to routers in the Service Group. This component includes the list of routers that have sent the web-cache a WCCP2_I_SEE_YOU message and a list of web-caches learnt from the WCCP2_I_SEE_YOU messages. The Web-Cache View Info Component also includes a change number which is incremented each time Service Group membership information changes.

3.7. Security

WCCP V2 provides a security component in each protocol message to allow simple authentication. Two options are currently supported:

* No security (default)
* MD5 password security

MD5 password security requires that each router and web-cache wishing to join a Service Group is configured with a matching Service Group password. Each WCCP protocol packet sent by a router or web-cache for that Service Group will contain in its security component the MD5 [RFC1321] checksum of the Service Group password and the WCCP protocol message (including the WCCP message header). Each web-cache or router in the Service Group will authenticate the security component in a received WCCP message immediately after validating the WCCP message header. Packets failing authentication, or lacking the expected authentication option, will be discarded.

3.8. Distribution of Traffic Assignments

WCCP V2 allows the traffic assignment method to be negotiated. There are two types of information to be communicated depending on the assignment method selected:

* Hash Tables
* Mask/Value Sets

3.8.1. Hash Tables

When using hash assignment each router uses a 256-bucket Redirection Hash Table to distribute traffic for a Service Group across the member web-caches. It is the responsibility of the Service Group's designated web-cache to assign each router's Redirection Hash Table.

The designated web-cache uses a WCCP2_REDIRECT_ASSIGNMENT message to assign the routers' Redirection Hash Tables. This message is generated following a change in Service Group membership and is sent to the same set of addresses to which the web-cache sends WCCP2_HERE_I_AM messages. The designated web-cache will wait for a time period of 1.5 * RA_TIMER_BASE_T following a membership change before generating the message in order to allow time for the Service Group membership to stabilise.

The designated web-cache lists the web-caches to which traffic should be distributed in either an Assignment Info Component or an Alternate Assignment Component within a WCCP2_REDIRECT_ASSIGNMENT message. Only those web-caches seen by every router in the Service Group are included.

The Assignment Info Component or Alternate Assignment Component within a WCCP2_REDIRECT_ASSIGNMENT message contains an Assignment Key. This will be reflected back to the designated web-cache in subsequent WCCP2_I_SEE_YOU messages from the routers in the Service Group. A WCCP2_REDIRECT_ASSIGNMENT message may be repeated after TRANSMIT_T time has elapsed if inspection of the Assignment Key within a WCCP2_I_SEE_YOU message indicates that a router has not received the assignment message.

A router will flush its Redirection Hash Table if a valid WCCP2_REDIRECT_ASSIGNMENT message has not been received within a time period of 5 * RA_TIMER_BASE_T following a Service Group membership change. To be valid, the message must contain the correct "Receive ID" and membership change number for the router.

Following successful receipt of a WCCP2_REDIRECT_ASSIGNMENT message, each router advertises its assigned Redirection Hash Table in all

subsequent WCCP2_HERE_I_AM messages. The Redirection Hash Table can be specified within an optional Alternate Assignment Map Component. If that component is not present, the current assignments for each web-cache are listed within the Web-Cache Identity Elements of the Router View Info Component.

3.8.2. Mask/Value Sets

When using mask assignment each router uses masks and a table of values to distribute traffic for a Service Group across the member web-caches. It is the responsibility of the Service Group's designated web-cache to assign each router's mask/value sets.

The designated web-cache uses a WCCP2_REDIRECT_ASSIGNMENT message to assign the routers' mask/value sets. This message is generated following a change in Service Group membership and is sent to the same set of addresses to which the web-cache sends WCCP2_HERE_I_AM messages. The designated web-cache will wait for a time period of 1.5 * RA_TIMER_BASE_T following a membership change before generating the message in order to allow time for the Service Group membership to stabilise.

The designated web-cache lists the web-caches to which traffic should be distributed in the Alternate Assignment Component of the WCCP2_REDIRECT_ASSIGNMENT message. Only those web-caches seen by every router in the Service Group are included.

The Alternate Assignment Component within a WCCP2_REDIRECT_ASSIGNMENT message contains an Assignment Key. This will be reflected back to the designated web-cache in subsequent WCCP2_I_SEE_YOU messages from the routers in the Service Group. A WCCP2_REDIRECT_ASSIGNMENT message may be repeated after TRANSMIT_T time has elapsed if inspection of the Assignment Key within a WCCP2_I_SEE_YOU message indicates that a router has not received the assignment message.

A router will flush its mask/value sets if a valid WCCP2_REDIRECT_ASSIGNMENT message has not been received within a time period of 5 * RA_TIMER_BASE_T following a Service Group membership change. To be valid, the message must contain the correct "Receive ID" and membership change number for the router.

Following successful receipt of a WCCP2_REDIRECT_ASSIGNMENT message, each router advertises its assigned mask/value sets in all subsequent WCCP2_HERE_I_AM messages. The mask/value sets can be listed within an optional Assignment Map Component or Alternate Assignment Map Component. If neither of those components is present, the current assignments for each web-cache are listed within the Web-Cache Identity Elements of the Router View Info Component.

3.9. Electing the Designated Web-cache

Election of the designated web-cache will take place once the Service Group membership has stabilised following a change. The designated web-cache must be receiving a WCCP2_I_SEE_YOU message from every router in the Service Group.

Election of the designated web-cache is not part of the WCCP protocol. However it is recommended that the eligible web-cache with the lowest IP address is selected as the designated web-cache for a Service Group.

3.10. Traffic Interception

A router will check packets passing through it against its set of Service Group descriptions. The Service Group descriptions are checked in priority order. A packet which matches a Service Group description is a candidate for redirection to a web-cache in the Service Group.

A router will not redirect a packet with a source IP address matching any web-cache in the Service Group.

3.11. Traffic Redirection

3.11.1. Redirection with Hash Assignment

To redirect a packet using hash assignment, a primary key is formed from the packet and hashed to yield an index into the Redirection Hash Table. The elements of the packet used to form the primary key are determined by the Service Group description.

If the indexed Redirection Hash Table entry is unassigned the packet is forwarded normally. If the entry contains only a web-cache index then the packet is redirected to that web-cache. Alternatively, if the entry is flagged as requiring an alternative hash then a secondary key is formed from the packet and hashed to yield a secondary index into the Redirection Hash Table. The elements of the packet used to form the secondary key are determined by the Service Group description.

If the secondary entry contains a web-cache index then the packet is redirected to that web-cache. If the secondary entry is unassigned the packet is forwarded normally. The alternative hashing flag in the secondary entry is ignored.

3.11.2. Redirection with Mask Assignment

To redirect a packet using mask assignment, a bitwise AND operation is performed between the mask from the first mask/value set assigned to the Service Group and the corresponding contents of the packet.

The masking operation is applied to both the source and destination IP addresses of the packet. For TCP and UDP packets, the masking operation is also applied to both the source and destination port numbers of the packet, when available. When port numbers are not available from a packet, the source and destination port elements of the result will be set to zero.

The output of this operation is compared against each entry in the list of value elements within the mask/value set. If a match is found the packet is redirected to the web-cache associated with the matching value element. If no match is found the process is repeated for each mask/value set defined for the Service Group. If no match is found after trying all of the mask/value sets defined for the Service Group, the packet is forwarded normally.

Mask/value sets are processed in the order in which they are presented in the Alternate Assignment Component. Similarly, value elements are compared in the order in which they are presented in a mask/value set.

3.12. Traffic Forwarding

WCCP V2 allows the negotiation of the forwarding method between a router and a web-cache (see Section 3.5.1). The currently defined forwarding methods are:

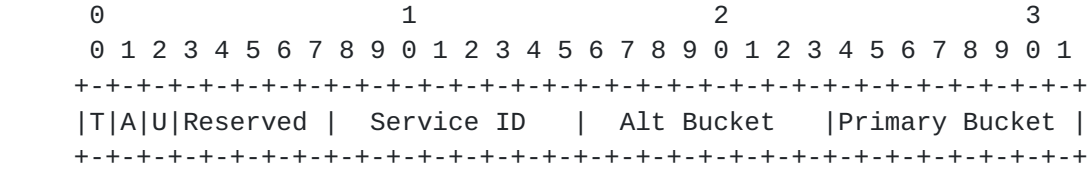
* GRE encapsulation
* Unencapsulated with L2 rewrite

3.12.1. Forwarding using GRE Encapsulation

Using this forwarding method, redirected packets are encapsulated in a new IP packet with a GRE [RFC1701] header followed by a 4-octet Redirect Header. The information provided within the Redirect Header can be used only if the U bit in the Redirect Header is 0. If the U bit is 1, the redirected packet is valid and should be processed normally, but the rest of the information within the 4-octet Redirect Header is unavailable and must be ignored.

The GRE encapsulation uses the simple 4-octet GRE header with the Flags and Version octets set to zero and a Protocol Type of 0x883E.

The Redirect Header is defined as follows:



T: Type of Service
   0: Standard (well known) service
   1: Dynamic service

A: Alternative bucket used (only valid for hash assignment)
   0: Primary bucket used
   1: Alternative bucket used

U: Unavailable
   0: Redirect Header contents are valid
   1: Redirect Header is present,
      but its contents (other than this bit) should
      be ignored and treated as being unavailable

Reserved

   Must be zero.

Service ID

Service Group identifier of the service that redirected this packet.

Alt Bucket

Alternative bucket index used to redirect the packet. Only valid for hash assignment.

Primary Bucket

Primary bucket index used to redirect the packet. Only valid for hash assignment.

3.12.2. Forwarding using L2 Rewrite

Using this forwarding method, redirected packets are not encapsulated. The router replaces the packet's destination MAC address with the MAC address of the target web-cache. The packet's source MAC address is set to the router's MAC address.

This forwarding method requires that the target web-cache is directly connected to the router at Layer 2. A router should not allow a web-cache to successfully negotiate this forwarding method unless it has been verified that the web-cache is directly connected.

A packet should not be redirected using this method if the packet's source MAC address matches the MAC address of a web-cache in the Service Group. See Section 3.13.3 for further details.

3.13. Packet Return

WCCP V2 allows a web-cache to decline a redirected packet and return it to the router for normal forwarding without further redirection. The method by which packets are returned from a web-cache to a router can be negotiated (see Section 3.5.3). The currently defined packet return methods are:

* GRE encapsulation
* Unencapsulated with L2 rewrite

3.13.1. Packet Return using GRE Encapsulation

Using this packet return method, a web-cache sends returned packets to a router using GRE encapsulation. Returned packets are encapsulated in a GRE packet [RFC1701] with a Protocol Type of 0x883E and containing either the Redirect Header from the originally redirected packet, or a Redirect Header with the U bit set if a valid Redirect Header was not present in the originally redirected packet. If the U bit is set, all other parts of the Redirect Header should be zero.

See Section 3.12.1 for the Redirect Header definition.

The receiving router removes the GRE encapsulation from each returned packet and forwards it without attempting further redirection.

3.13.2. Packet Return using L2 Rewrite

Using this packet return method, returned packets are not encapsulated, so any encapsulation added by the router during redirection must be removed by the web-cache. The web-cache then replaces the packet's destination MAC address with the router's MAC address and sets the packet's source MAC address to the web-cache's own MAC address.

The packet return method requires that the router receiving the return packet does not attempt to redirect it again, otherwise the packet will repeatedly loop between the router and the web-cache.

3.13.3. Preventing redirection of returned packets

When a router receives a returned packet it must not attempt to redirect the packet back to a web-cache. Three methods are available to prevent further redirection:

* Encapsulation
* Source MAC address check
* Interface configuration

The encapsulation method requires a web-cache to send returned packets to a router using GRE encapsulation, as described in Section 3.13.1. Returned packets are identified using the web-cache's source IP address and/or the GRE Protocol Type of 0x883E. Following removal of the GRE encapsulation these packets must be excluded from further redirection.

The source MAC address check method requires a web-cache to return a packet unencapsulated to the router using L2 rewrite, as described in Section 3.13.2. The router must record the MAC address of each web-cache that has successfully negotiated the L2 rewrite packet return method. The router then excludes from redirection any packet received with a source MAC address belonging to one of the known web-caches.

The interface configuration method requires that a router is configured to inhibit redirection of packets arriving on an interface connected to one or more web-caches. The suitability of this mechanism is dependant on the network topology. It is only required if the source MAC address check cannot be used in combination with the L2 rewrite return method.

3.14. Querying Web-Cache Time-Out

If a router does not receive a WCCP2_HERE_I_AM message from a Service Group member during a time period of 2.5 * TIMEOUT_BASE_T it will query the member by sending a unicast WCCP2_REMOVAL_QUERY message to it. The target Service Group member should respond by sending a series of three identical unicast WCCP2_HERE_I_AM messages to the router, each separated by a time interval of 0.1 * TRANSMIT_T.

If a router does not receive a WCCP2_HERE_I_AM message from a Service Group member during a time period of 3 * TIMEOUT_BASE_T it will consider the member to be unusable and remove it from the Service Group. The web-cache will no longer appear in the Router View Info Component of the WCCP2_I_SEE_YOU message. The web-cache will also be purged from the assignment data for the Service Group.

3.15. Sending additional WCCP2_HERE_I_AM messages

If a web-cache does not receive a WCCP2_I_SEE_YOU message from a router in response to a unicast WCCP2_HERE_I_AM message after a time period of 0.5 * TRANSMIT_T has elapsed, the web-cache may optionally choose to transmit a new WCCP2_HERE_I_AM message at this moment instead of waiting for a full TRANSMIT_T time interval to elapse.

This action is permitted only if, in response to the previous WCCP2_HERE_I_AM message unicast to the router, the web-cache successfully received a WCCP2_I_SEE_YOU message from the router in which the web-cache appeared in the Router View Info Component of the message.

The web-cache may continue transmitting WCCP2_HERE_I_AM messages at time intervals of 0.5 * TRANSMIT_T until a WCCP2_I_SEE_YOU message is received from the router, or until a total of 6 WCCP2_HERE_I_AM messages have been transmitted since the last WCCP2_I_SEE_YOU message was received.

3.16. Command and Status Information

WCCP V2 includes a mechanism to allow web-caches to send commands to routers within a service group. The same mechanism can be used by the routers to provide status information to web-caches.

The mechanism is implemented by the Command Extension Component. This component is included in the WCCP2_HERE_I_AM message from a web-cache passing commands to routers in a Service Group.

If a router needs to send status information back to a web-cache it will include a command in the Command Extension Component within its own WCCP2_I_SEE_YOU message. That command will indicate the type of status information being carried.

4. Protocol Messages

4.1. Overview

Each WCCP protocol message is carried within a UDP packet with source and destination ports of 2048. Every WCCP message begins with a fixed-length 8-octet header, followed by a number of additional variable-length components.

The WCCP header specifies the message type, the major and minor protocol version numbers, and the length of the remainder of the message. Any contents of the UDP packet extending beyond this specified message length must be ignored.

There are four WCCP V2 message types:

* Here I Am
* I See You
* Redirect Assign
* Removal Query

Messages with a header containing an unrecognised type or the incorrect major version number must be ignored. Note that messages containing the correct major version number but an unrecognised minor version number should continue to be processed.

Every component following the WCCP header conforms to a Type-Length-Value (TLV) format. Each component begins with a 2-octet type followed by a 2-octet length. The length specifies the number of octets remaining within the component following the length field. The specified length must be a multiple of 4 octets. Padding is allowed within each component, but no padding is allowed between components, therefore the length of a component must correctly specify the offset to the beginning of the subsequent component.

The type of a component specifies the format of the data it contains. If the component type is not recognised by the receiver, the number of following octets specified in the length field must be ignored and message processing should resume at the beginning of the next component.

Some components contain nested elements which also conform to a TLV format. In general, when the type of a nested TLV element is unrecognised, only the smallest unrecognised element should be ignored.

If the length of a component extends beyond the end of the WCCP message (as specified in the WCCP header), the whole component must

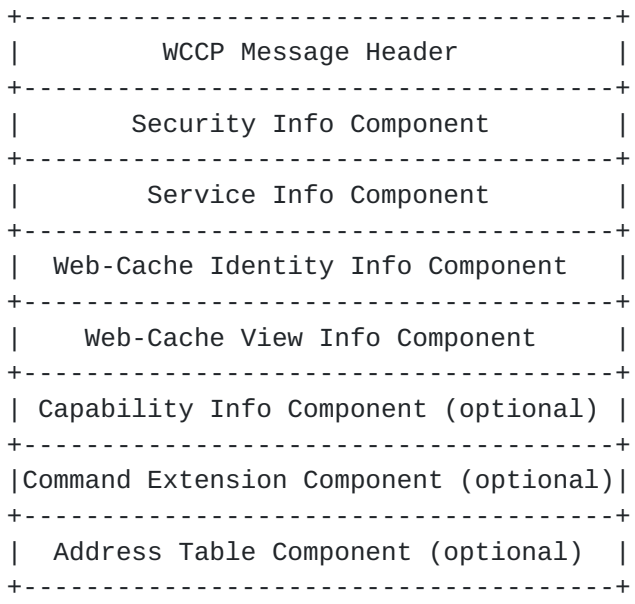
be ignored.

If a message contains multiple components of the same type and only a single component of that type is expected, the first element of that type should be processed normally and any subsequent elements of the same type should be ignored.

In general, receivers should be tolerant of unexpected components and elements within a message, being mindful of the fact that the protocol is extensible. Protocol extensions may be added with or without a minor version increment, depending on the nature of the extension.

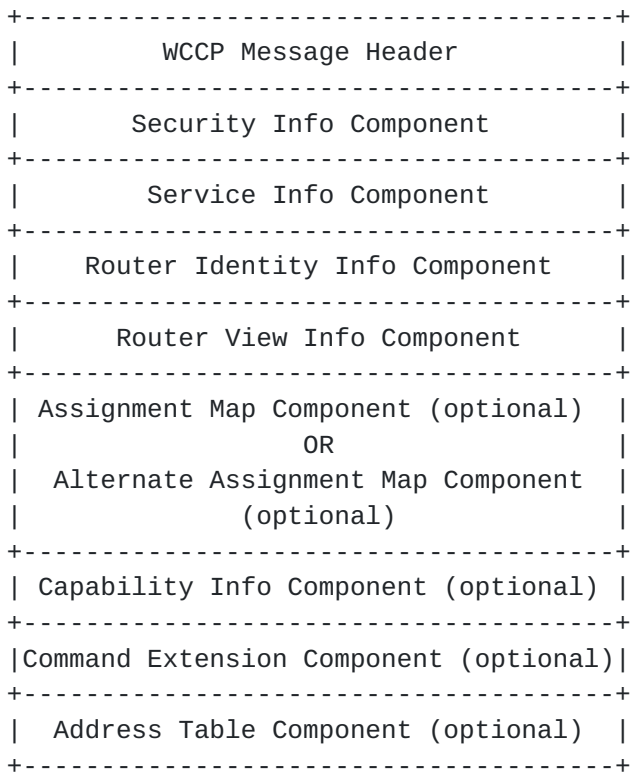
4.2. 'Here I Am' Message

A 'Here I Am' message contains the following components:



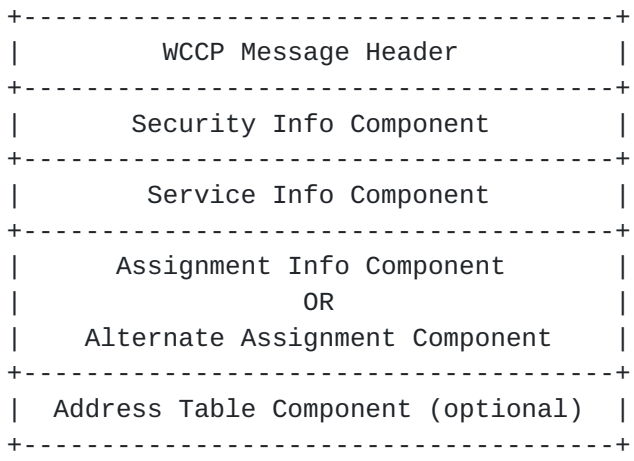
4.3. 'I See You' Message

An 'I See You' message contains the following components:



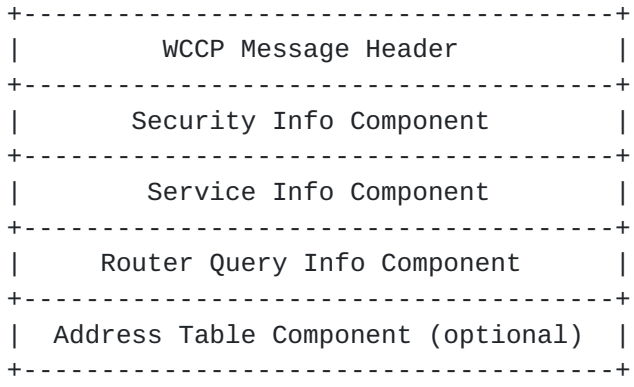
4.4. 'Redirect Assign' Message

A 'Redirect Assign' message contains the following components:

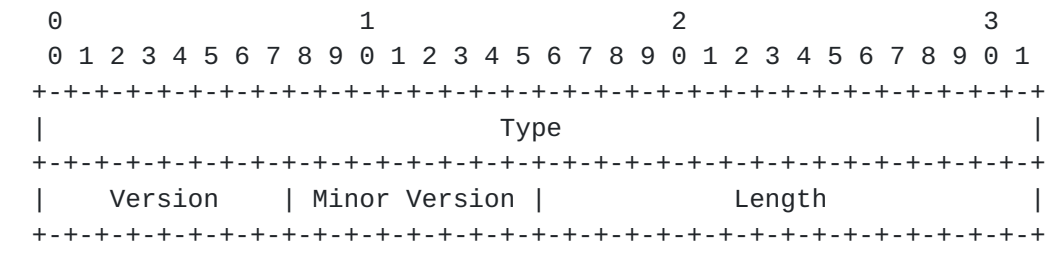


4.5. 'Removal Query' Message

A 'Removal Query' message contains the following components:



4.6. WCCP Message Header



Type

Indicates the type of the WCCP message. The following types are defined:

0x0A - WCCP2_HERE_I_AM (10)
0x0B - WCCP2_I_SEE_YOU (11)
0x0C - WCCP2_REDIRECT_ASSIGN (12)
0x0D - WCCP2_REMOVAL_QUERY (13)

Version

Indicates the protocol version required to process the message. The value defined by this document is:

0x02 - WCCP V2

Minor Version

Indicates a minor revision level of the protocol that the sender supports and which the message conforms to. The use of different protocol revision levels is described in Section 3.4. The values defined by the current revision of this document are:

0x00 - Protocol Version 2.00
0x01 - Protocol Version 2.01

Length

Length of the WCCP message not including the WCCP Message Header.

4.7. Multiple Address family support

By default, network addresses used within the protocol are IPv4 addresses. However, with protocol version 2.01, alternative address families can be used whenever the optional address table component is present in a protocol message.

All addresses and address masks used within a protocol message are referenced via a 4-octet address element. This element can contain:

* the special value of 0 indicating an unspecified address, or
* an IPv4 address or mask, or
* the value of an address index.

The address index is an indirect reference to an address or mask entry within the address table component which is contained within the same protocol message. Address indices are numbered from 1 upwards.

If an address table component is present in a message, every address element within the message contains either an address index or an unspecified address.

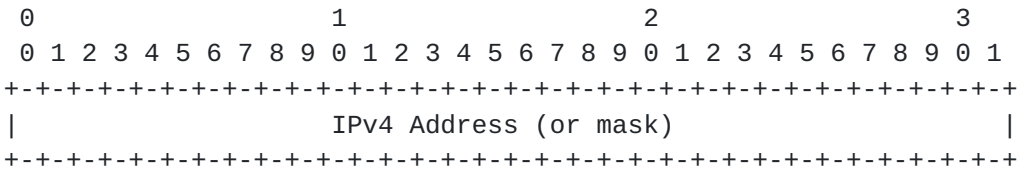
When a WCCP message has a protocol version of 2.01, the correct interpretation of each non-zero address element requires knowledge of the presence of an address table component. Therefore, there is a requirement to check for the existence of an address table component before attempting to interpret any non-zero address elements within the message.

If an address table component is not present in a message, every address element within the message contains an IPv4 address or mask. Address tables are not permitted when the protocol version is 2.00.

4.7.1. Messages without an address table component

When an address table component is not present, every network address (or mask) within the protocol message is specified as follows:

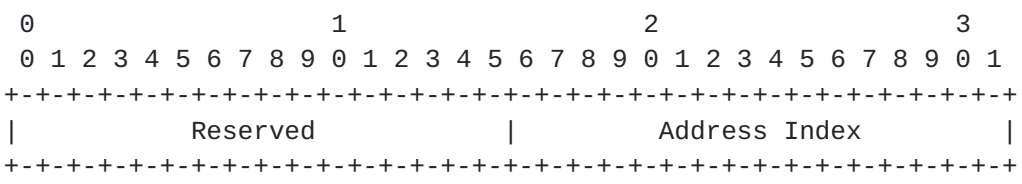
Address Element:



4.7.2. Messages with an address table component

When an address table component is present in a protocol message, every address element within the same message is specified as follows:

Address Element:



Reserved

Must be zero.

Address Index

An index into the list of network addresses provided in the address table component defined in Section 5.1.5. The first address in the table is referenced using index 1, the second address is referenced using index 2, and so on. Address indices that would fall beyond the length of the address table component are invalid. A value of 0 is special and will be interpreted as an unspecified address (or an address mask with no bits set).

5. Message Components

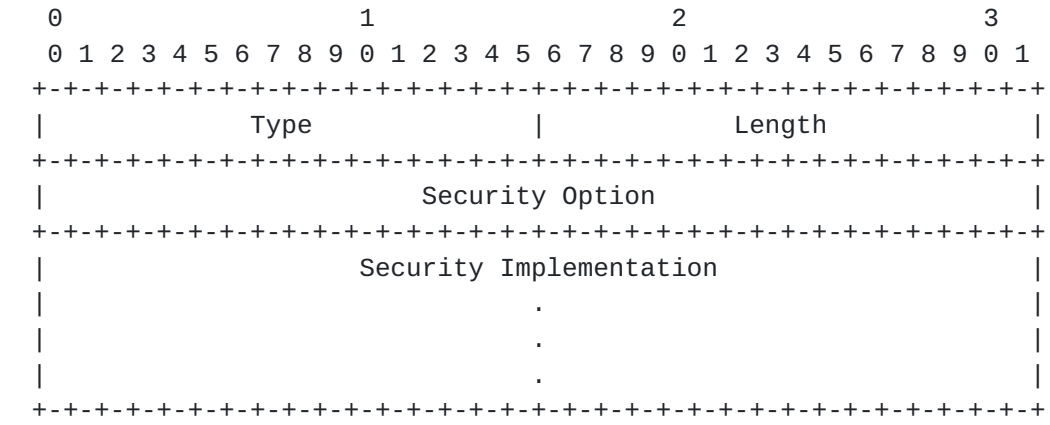
Each WCCP message comprises a WCCP Message Header followed by a number of message components, some of which have a variable length. The defined components are:

* Security Info
* Service Info
* Capabilities Info
* Command Extension
* Address Table
* Web-Cache Identify Info
* Web-Cache View Info
* Router Identity Info
* Router View Info
* Assignment Map
* Alternate Assignment Map
* Assignment Info
* Alternate Assignment
* Router Query Info

Note that components are padded to align on a 4-octet boundary. Each component has a 4-octet header specifying the component type and length. The length value does not include the 4-octet component header.

5.1. Components used in multiple message types

5.1.1. Security Info Component



Type

0x00 - WCCP2_SECURITY_INFO (0)

Length

Length of the remainder of the component.

Security Option

The currently defined values are:

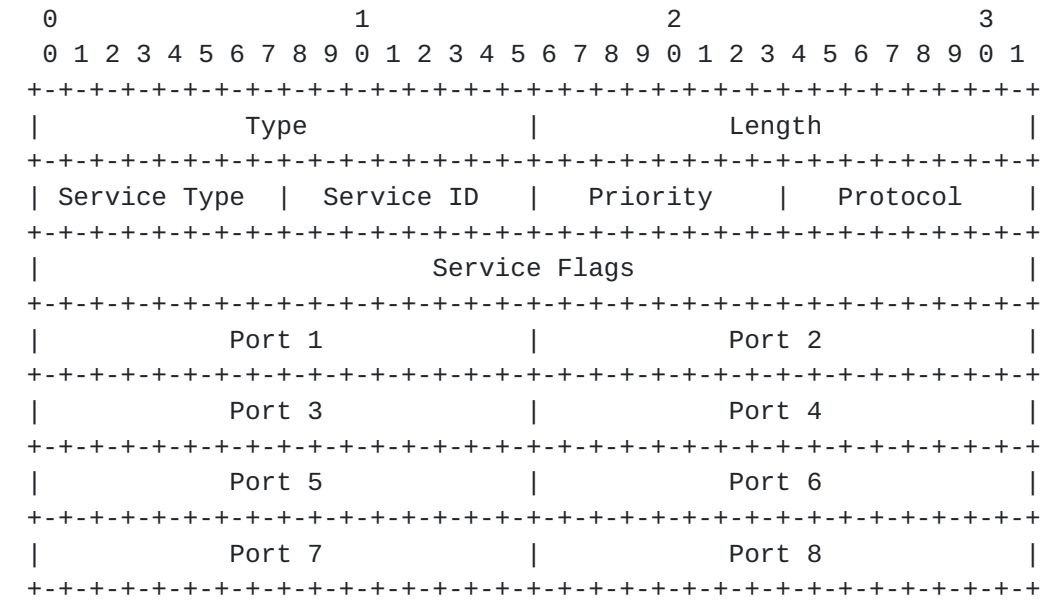
0x00 - WCCP2_NO_SECURITY
0x01 - WCCP2_MD5_SECURITY

Security Implementation

If Security Option has the value WCCP2_NO_SECURITY this field is not present. If Security Option has the value WCCP2_MD5_SECURITY this is a 16-octet field containing the MD5 [RFC1321] checksum of the WCCP message and the Service Group password. The maximum password length is 8 octets.

Prior to calculating the MD5 checksum the password should be padded out to 8 octets with trailing zeros and the Security Implementation field of the Security Option set to zero. The MD5 checksum is calculated using the 8-octet padded password followed by the WCCP message (including the WCCP Message Header).

5.1.2. Service Info Component



Type

0x01 - WCCP2_SERVICE_INFO (1)

Length

Length of the remainder of the component.

Service Type

The following service types are currently defined:

0x00 - WCCP2_SERVICE_STANDARD

The service is a well known service and is described by the Service ID. All service definition fields other than Service ID should be zero.

0x01 - WCCP2_SERVICE_DYNAMIC

The service is a dynamic service as is defined by the Protocol, Service Flags and Port fields.

Service ID

The service number which, in combination with the service type, uniquely identifies the service group. For services of type WCCP2_SERVICE_DYNAMIC, all values from 0 to 255 inclusive are valid. For services of type WCCP2_SERVICE_STANDARD, a single service number is currently defined:

0x00 - HTTP (Protocol: TCP, Destination Port: 80)

Priority

Service priority. The lowest priority is 0, the highest is 255. Packets for redirection are matched against Services in priority order, highest first. Well known services have a priority of 240.

Protocol

IP protocol identifier. The protocol type of traffic to be redirected. A value of 0 indicates that all protocol types should be redirected, unless the "Redirect Only Protocol 0" flag is set (in which case only protocol 0 would be redirected).

Service Flags

0x0001 - Source IP Hash
0x0002 - Destination IP Hash
0x0004 - Source Port Hash
0x0008 - Destination Port Hash
0x0010 - Ports Defined
0x0020 - Ports Source
0x0040 - Redirect Only Protocol 0 (* see note)
0x0100 - Source IP Alternative Hash
0x0200 - Destination IP Alternative Hash
0x0400 - Source Port Alternative Hash
0x0800 - Destination Port Alternative Hash

(* - requires minimum protocol version 2.01)

The primary hash flags (Source IP Hash, Destination IP Hash, Source Port Hash, Destination Port Hash) determine which protocol header fields of a packet will be hashed to yield the Redirection Hash Table primary bucket index. The hash index is constructed by XORing each octet of the appropriate fields from the packet header. The hash index is a single octet and has an initial value of zero.

If alternative hashing has been enabled for the primary bucket (see the bucket definition in Section 6.5), the alternate hash flags (Source IP Alternative Hash, Destination IP Alternative Hash, Source Port Alternative Hash, Destination Port Alternative Hash) determine which protocol header fields of a packet will be hashed to yield a secondary bucket index. The secondary hash index is constructed by XORing each octet of the appropriate fields from the packet header. The secondary hash index is a single octet and has an initial value of zero.

The primary hash flags and alternate hash flags are valid only when the service group uses hash assignment, in which case at least one primary hash flag and one secondary hash flag must be set.

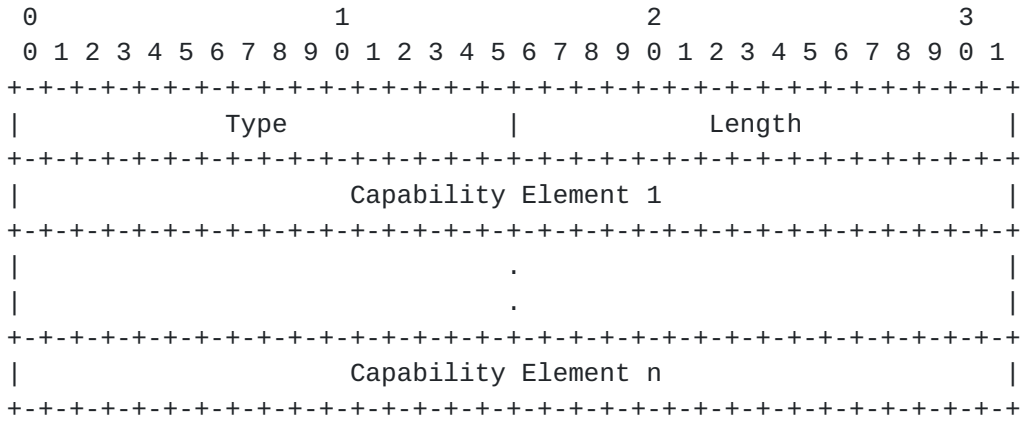
Port 1 -> Port 8

A list of UDP or TCP port numbers. The port list is active only if the service protocol is set to UDP or TCP and the service flag "Ports Defined" is set.

If the "Ports Source" flag is set the port information refers to the source port within a packet to be redirected, if clear the port information refers to the destination port within a packet to be redirected. When the list is active, a packet can be redirected only if it uses one of the port numbers contained in this list.

If less than eight ports are specified, the list is terminated with a port value of zero, in which case subsequent entries in the list are ignored.

5.1.3. Capabilities Info Component



Type

0x08 - WCCP2_CAPABILITY_INFO (8)

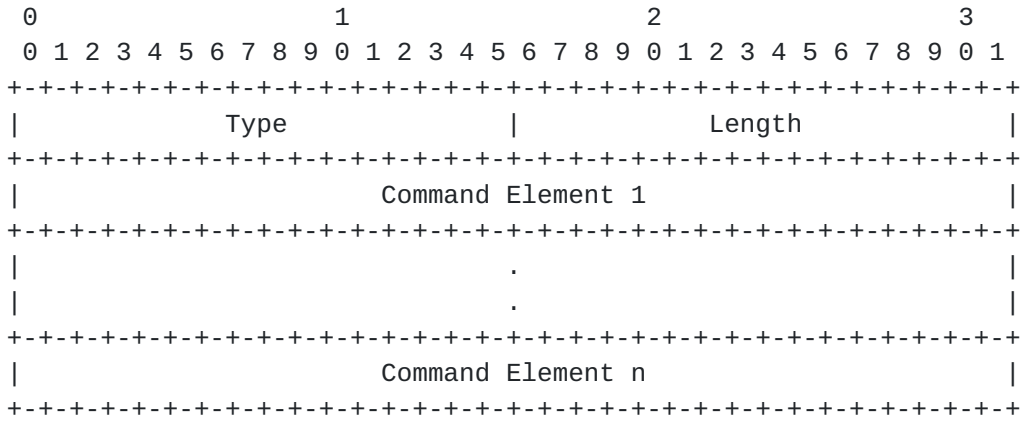
Length

Length of the remainder of the component.

Capability Element 1 -> Capability Element n

Elements in TLV-format each describing a router or web-cache capability. Each element is defined in Section 6.11.

5.1.4. Command Extension Component



Type

0x0F - WCCP2_COMMAND_EXTENSION (15)

Length

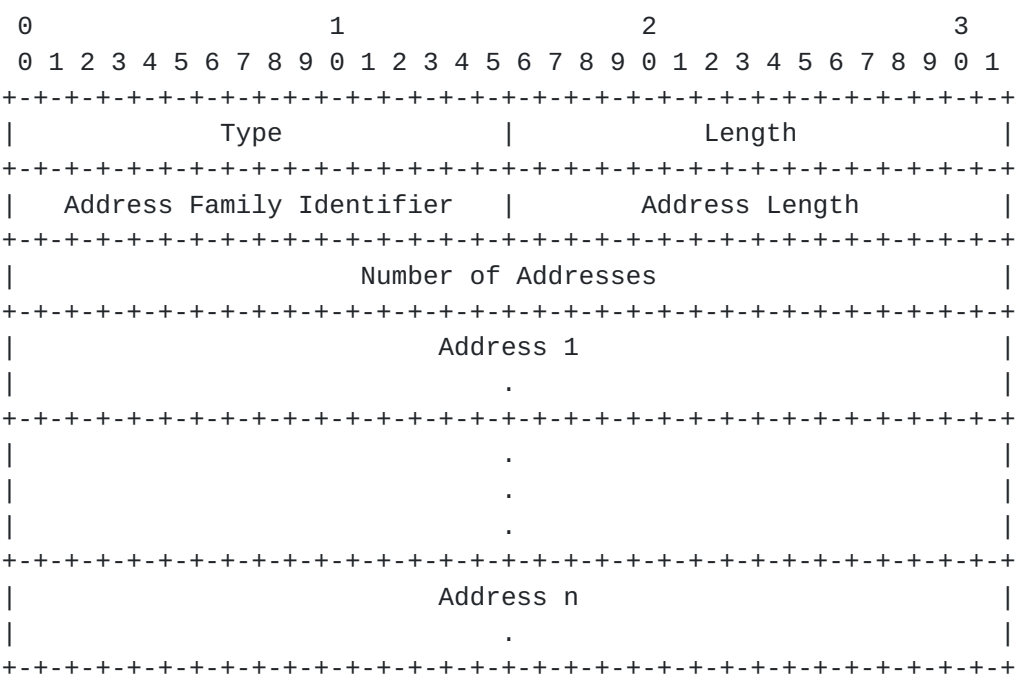
Length of the remainder of the component.

Command Element 1 -> Command Element n

Elements in TLV-format each containing a router or web-cache command. Each element is defined in Section 6.12.

5.1.5. Address Table Component

This component is valid from protocol version 2.01. It provides a list of network addresses that are referenced within the WCCP message. References to these addresses are made via address elements within other WCCP message components. The referencing address element is defined in Section 4.7.2.



Type

0x11 - WCCP2_ADDRESS_TABLE (17)

Length

Length of the remainder of the component.

Address Family Identifier

Indicates the address family of all network addresses within the table. The values are defined by the Internet Assigned Numbers Authority (IANA) Address Family Numbers registry [IANA-AF]. Relevant values include:

0x02 - IP version 6 (IPv6)

As IPv4 addresses can be specified directly within a WCCP message without requiring an address table, the use of an IPv4 address table is unnecessary and therefore strongly discouraged.

Address Length

The length in octets of each entry within the list of network addresses. The length of each entry must be a multiple of 4 octets. If this length is larger than the natural size of an address of the given address family, excess trailing octets in each entry should be set to zero by the sender and ignored by the receiver.

Number of Addresses

The number of addresses (n) contained within the following list.

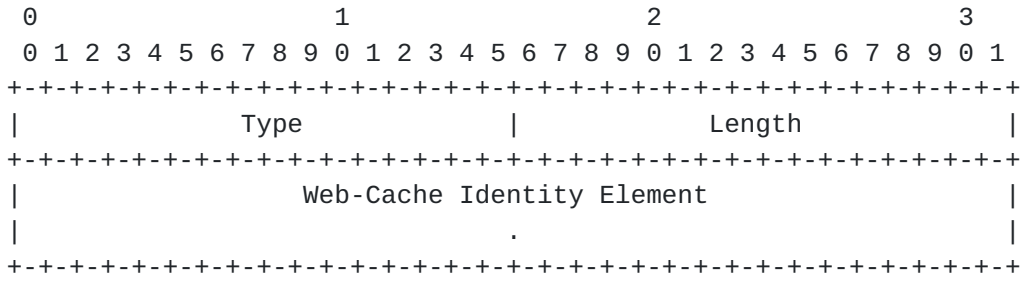
Address 1 -> Address n

A list of network addresses that can be referenced via their index in this list. The first address is referenced using index 1 and the last address is referenced using index n, providing a list of n addresses.

5.2. 'Here I Am' message components

The following sub-sections describe components used only in 'Here I Am' messages.

5.2.1. Web-Cache Identity Info Component



Type

0x03 - WCCP2_WC_ID_INFO (3)

Length

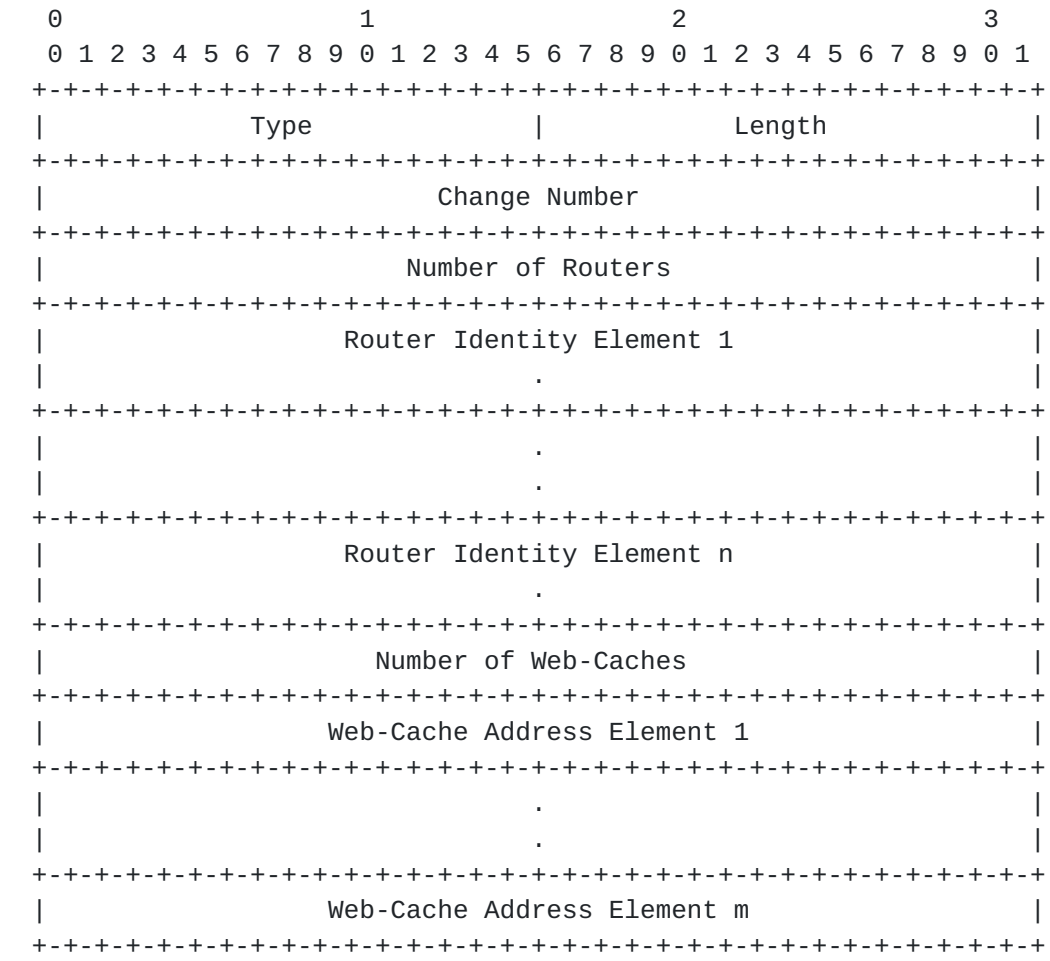
Length of the remainder of the component.

Web-Cache Identity Element

An element indicating the web-cache IP address and its redirection assignments. The element is defined in Section 6.4.

5.2.2. Web-Cache View Info Component

This component represents a web-cache's view of the Service Group.



Type

0x05 - WCCP2_WC_VIEW_INFO (5)

Length

Length of the remainder of the component.

Change Number

A value incremented each time there is a change in the view.

Number of Routers

The number of routers (n) in the Service Group.

Router Identity Element 1 -> Router Identity Element n

Elements indicating the identifying IP address for each router in the Service Group and the last "Receive ID" obtained from each. Each element is defined in Section 6.1.

Number of Web-Caches

The number of web-caches (m) in the Service Group.

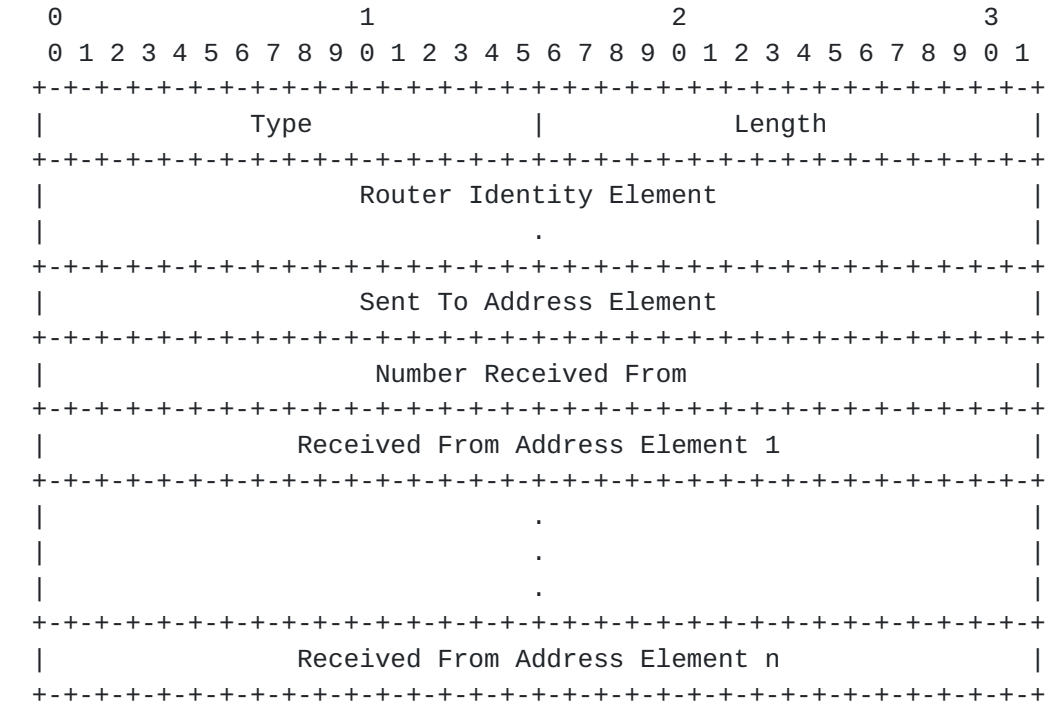
Web-Cache Address Element 1 -> Web-Cache Address Element m

Elements indicating the web-cache IP addresses learnt from WCCP2_I_SEE_YOU messages. Each address element is defined in Section 4.7.

5.3. 'I See You' message components

The following sub-sections describe components used only in 'I See You' messages.

5.3.1. Router Identity Info Component



Type

0x02 - WCCP2_ROUTER_ID_INFO (2)

Length

Length of the remainder of the component.

Router Identity Element

Element indicating the router's identifying IP address and "Receive ID". The identifying IP address must be a valid, reachable address for the router. The element is defined in Section 6.1.

Sent To Address Element

Identifies the IP address to which the target web-cache sent the WCCP2_HERE_I_AM message. When this component is present in a unicast WCCP2_I_SEE_YOU message, this element identifies the IP address that the target web-cache used. When present in a multicast WCCP2_I_SEE_YOU message, this element identifies the Service Group multicast address. The address element is defined in Section 4.7.

Number Received From

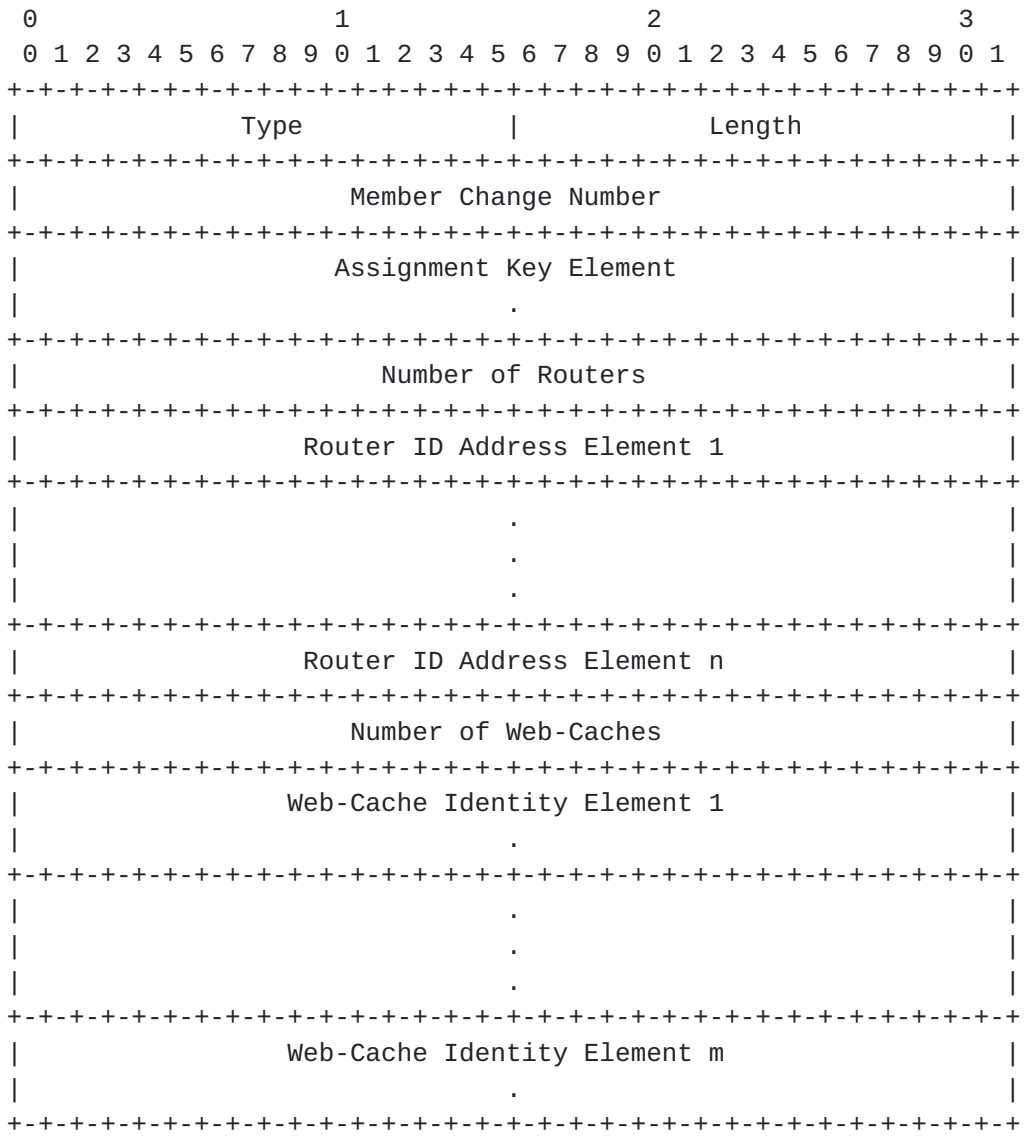
The number of web-caches (n) to which this message is directed. When using multicast addressing it may be less than the number of web-caches which actually see the message.

Received From Address Element 1 -> Received From Address Element n

Elements identifying the IP addresses of web-caches to which this message is directed. When using multicast addressing it may be a subset of the web-caches which actually see the message. Each address element is defined in Section 4.7.

5.3.2. Router View Info Component

This component represents a router's view of the Service Group.



Type

0x04 - WCCP2_RTR_VIEW_INFO (4)

Length

Length of the remainder of the component.

Member Change Number

A value incremented each time there is a change in the Service Group membership.

Assignment Key Element

The Assignment Key Element received in the most recent valid WCCP2_REDIRECT_ASSIGNMENT message. This is used by the designated web-cache to verify that an assignment has been accepted by the router and that the assignment remains active. The element is defined in Section 6.3.

Number of Routers

The number of routers (n) in the Service Group.

Router ID Address Element 1 -> Router ID Address Element n

Elements identifying the Router IDs of routers in the Service Group. The list is constructed from routers reported by web-caches via WCCP2_HERE_I_AM messages. Note that a router does not include itself in the list unless it has also been reported via a WCCP2_HERE_I_AM message. Each element is defined in Section 4.7.

Number of Web-Caches

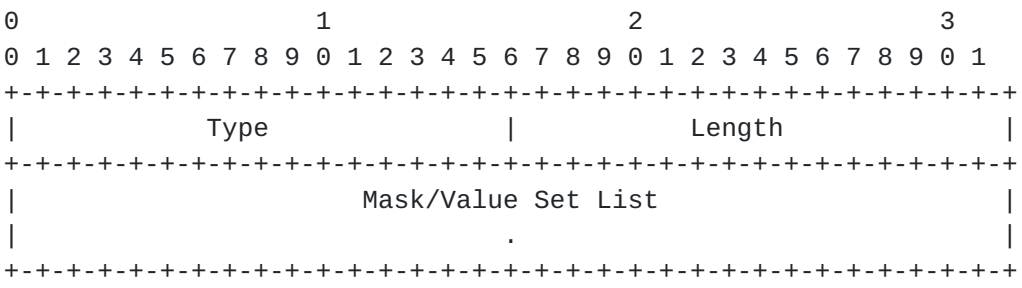
The number of useable web-caches (m) in the Service Group.

Web-Cache Identity Element 1 -> Web-Cache Identity Element m

Web-Cache Identity Elements of the useable web-caches in the Service Group. This list contains web-caches that have sent the router a WCCP2_HERE_I_AM message with a valid "Receive ID" and compatible capabilities. Each element is defined in Section 6.4.

5.3.3. Assignment Map Component

This component can only be used with Service Groups using mask assignment.



Type

0x0E - WCCP2_ASSIGNMENT_MAP (14)

Length

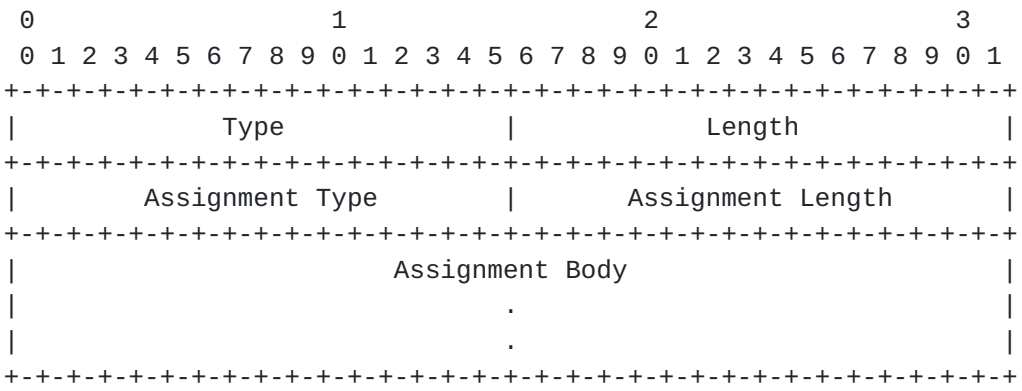
Length of the remainder of the component.

Mask/Value Set List

A list of mask/value sets. The list is defined in Section 6.13.

5.3.4. Alternate Assignment Map Component

This component is valid from protocol version 2.01.



Type

0x10 - WCCP2_ALT_ASSIGNMENT_MAP (16)

Length

Length of the remainder of the component.

Assignment Type

Indicates the format of Assignment Body. The currently defined values are:

0x00 - WCCP2_HASH_ASSIGNMENT
0x01 - WCCP2_MASK_ASSIGNMENT
0x02 - WCCP2_ALT_MASK_ASSIGNMENT

Assignment Length

Length of the remainder of the component (Assignment Body).

Assignment Body

The format of Assignment Body is specified by the value of Assignment Type, as follows:

WCCP2_HASH_ASSIGNMENT:

Hash Buckets Assignment Element (Section 6.5)

WCCP2_MASK_ASSIGNMENT:

Mask/Value Set List (Section 6.13)

WCCP2_ALT_MASK_ASSIGNMENT:

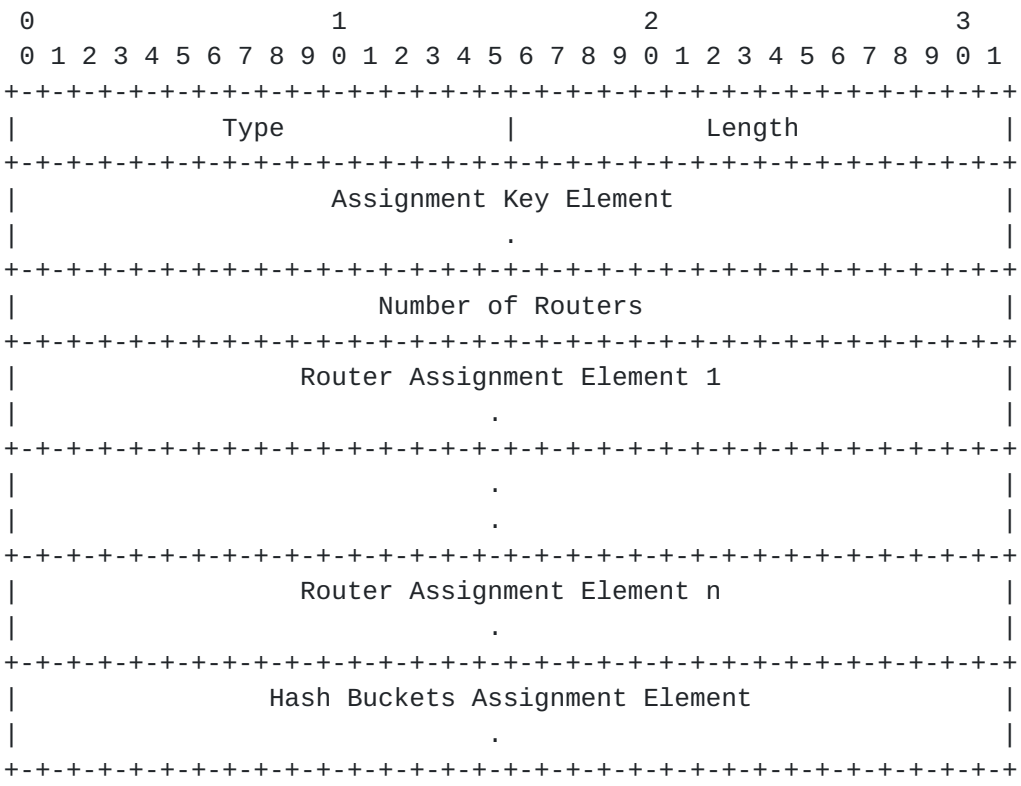
Alternate Mask/Value Set List (Section 6.17)

5.4. 'Redirect Assign' message components

The following sub-sections describe components used only in 'Redirect Assign' messages.

5.4.1. Assignment Info Component

This component can only be used with Service Groups using hash assignment.



Type

0x06 - WCCP2_REDIRECT_ASSIGNMENT (6)

Length

Length of the remainder of the component.

Assignment Key Element

The designated web-cache expects this element to be returned by a router in subsequent WCCP2_I_SEE_YOU messages. The element is defined in Section 6.3.

Number of Routers

The number of routers (n) reachable by the designated web-cache.

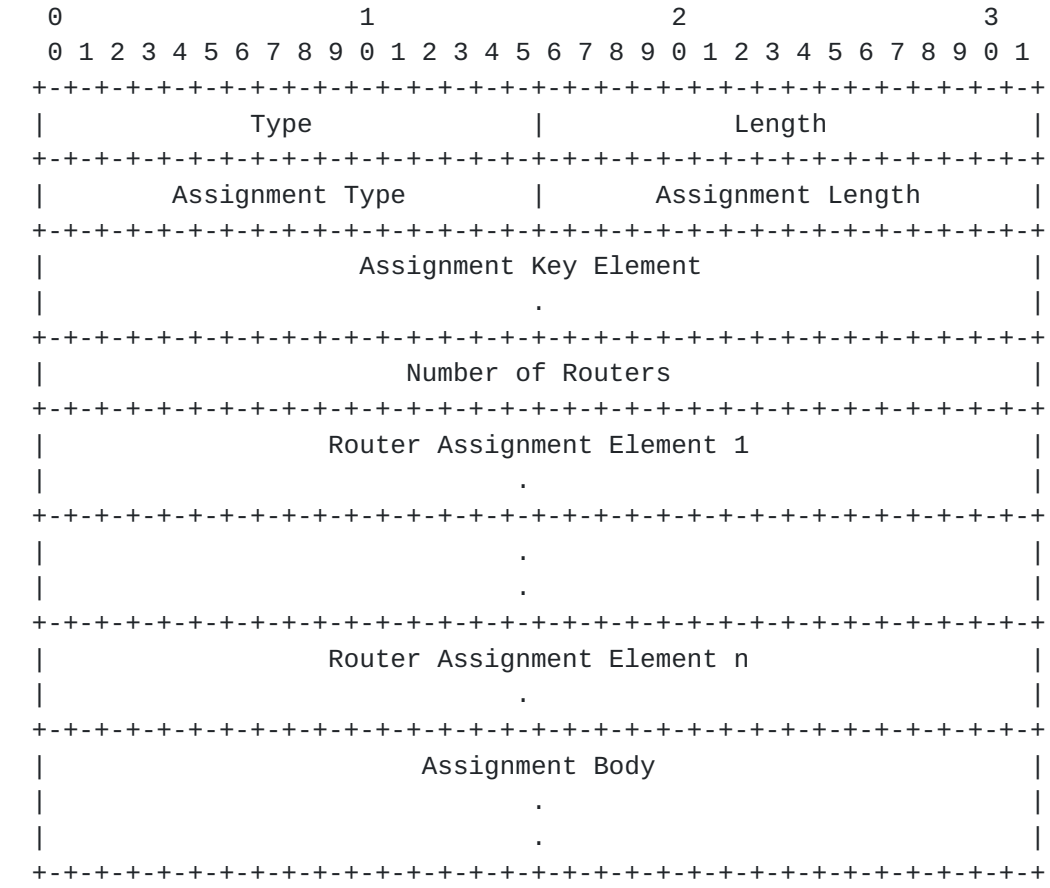
Router Assignment Element 1 -> Router Assignment Element n

Elements indicating the identifying IP address, "Receive ID" and "Change Number" for each router. Each element is defined in Section 6.2.

Hash Buckets Assignment Element

A list of web-caches and hash bucket assignments. The element is defined in Section 6.5.

5.4.2. Alternate Assignment Component



Type

0x0D - WCCP2_ALT_ASSIGNMENT (13)

Length

Length of the remainder of the component.

Assignment Type

Indicates the format of Assignment Body. The currently defined values are:

0x00 - WCCP2_HASH_ASSIGNMENT
0x01 - WCCP2_MASK_ASSIGNMENT
0x02 - WCCP2_ALT_MASK_ASSIGNMENT (* see note)

(* - requires minimum protocol version 2.01)

Assignment Length

Length of the remainder of the component (from Assignment Key Element onwards).

Assignment Key Element

The designated web-cache expects this element to be returned by a router in subsequent WCCP2_I_SEE_YOU messages. The element is defined in Section 6.3.

Number of Routers

The number of routers (n) reachable by the designated web-cache.

Router Assignment Element 1 -> Router Assignment Element n

Elements indicating the router ID address, "Receive ID" and "Change Number" for each router. Each element is defined in Section 6.2.

Assignment Body

The format of Assignment Body is specified by the value of Assignment Type, as follows:

WCCP2_HASH_ASSIGNMENT:

Hash Buckets Assignment Element (Section 6.5)

WCCP2_MASK_ASSIGNMENT:

Mask/Value Set List (Section 6.13)

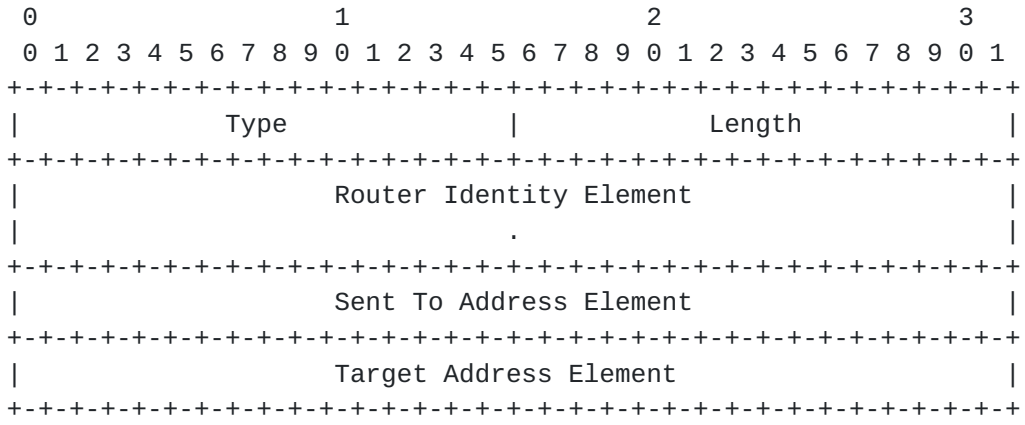
WCCP2_ALT_MASK_ASSIGNMENT:

Alternate Mask/Value Set List (Section 6.17)

5.5. 'Removal Query' message components

The following sub-section describes a component used only in 'Removal Query' messages.

5.5.1. Router Query Info Component



Type

0x07 - WCCP2_QUERY_INFO (7)

Length

Length of the remainder of the component.

Router Identity Element

Element indicating the router's identifying IP address and "Receive ID". The identifying IP address must be a valid, reachable address for the router. The element is defined in Section 6.1.

Sent To Address Element

Indicates the IP address to which the target web-cache sent its last received WCCP2_HERE_I_AM message. This will be the multicast address if the web-cache is multicasting its WCCP2_HERE_I_AM messages. The address element is defined in Section 4.7.

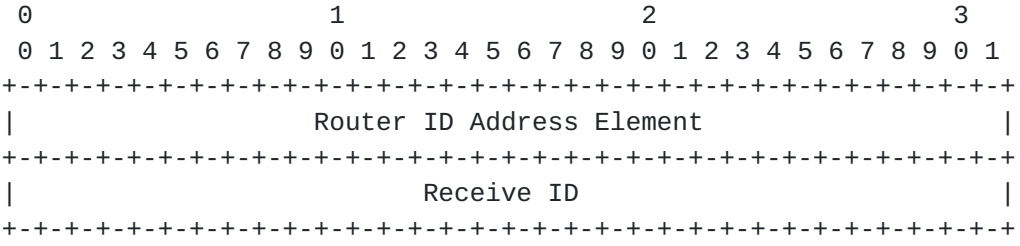
Target Address Element

Indicates the identifying IP address of the web-cache being queried. The address element is defined in Section 4.7.

6. Message Elements

The following sub-sections describe the message elements used within WCCP message components.

6.1. Router Identity Element



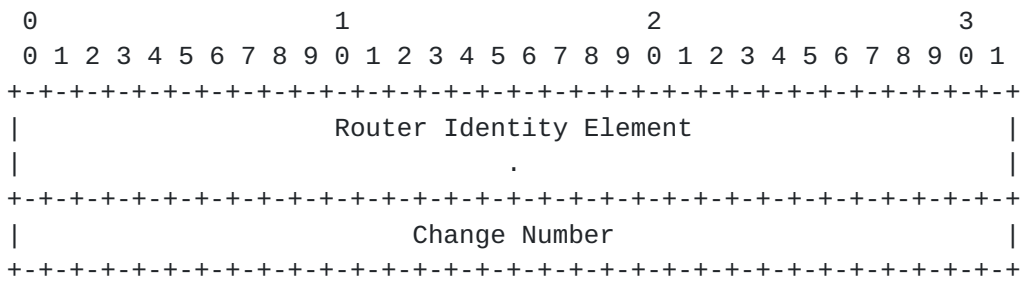
Router ID Address Element

Indicates the router's identifying IP address. The identifying IP address must be a valid IP address by which the router is reachable. The address element is defined in Section 4.7.

Receive ID

A number maintained by the router for each Service Group. It is incremented each time the router sends a WCCP protocol message that includes a Router Identity Info Component. A router's Receive ID will never be zero.

6.2. Router Assignment Element



Router Identity Element

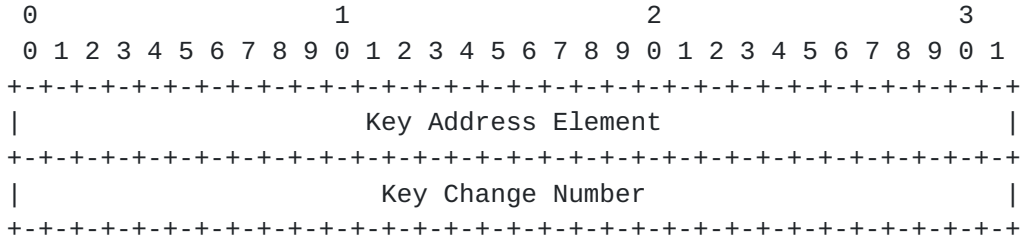
Indicates the router's identifying IP address and the last Receive ID obtained from it. The element is defined in Section 6.1. A router will ignore an assignment if the Receive ID is invalid.

Change Number

Last Member Change Number received from the router identified by the Router Identity Element. A router will ignore an assignment if Change Number is invalid.

6.3. Assignment Key Element

This element uniquely identifies a particular assignment.



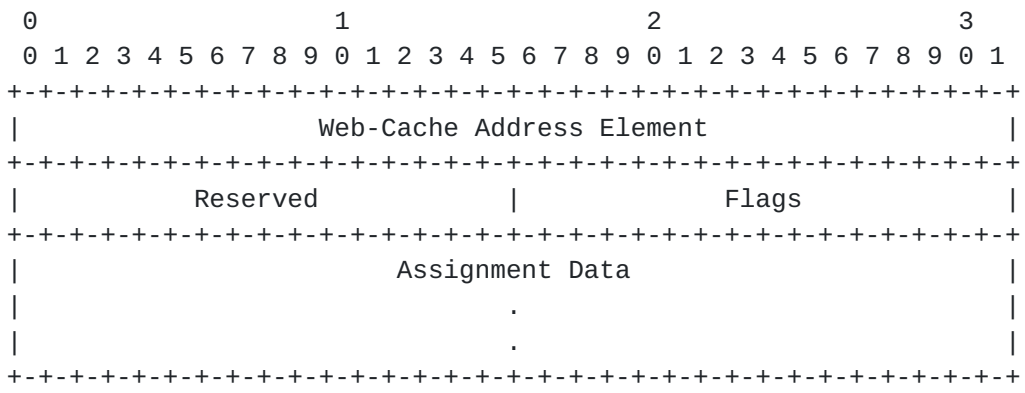
Key Address Element

Indicates the identifying IP address of the designated web-cache. The address element is defined in Section 4.7.

Key Change Number

A number maintained by the designated web-cache. It is incremented by the designated web-cache each time a change is made to the assignments for a Service Group.

6.4. Web-Cache Identity Element



Web-Cache Address Element

Indicates the identifying IP address of the web-cache. This must be a valid IP address by which the web-cache is reachable. The address element is defined in Section 4.7.

Reserved

Must be zero.

Flags

Bit 0 (U bit):

If set, this bit indicates that the web-cache does not have an assignment in the current Service Group assignments and that the assignment data which follows is historical. Historical data may be used by the designated web-cache to re-assign the same assignment entries to a web-cache that left and subsequently rejoined a Service Group.

Bit 1 & bit 2 (Type bits):

Two bits indicating the format of the Assignment Data element immediately following. The meaning of the bit settings are shown in the following table:

Bit 1	Bit 2	Meaning	
0	0	Hash Assignment	
1	0	Mask Assignment	
0	1	No Assignment	(* see note)
1	1	Extended Assignment	(* see note)

(* - requires minimum protocol version 2.01)

Bit 3 (V bit):

If set, this bit indicates that the protocol version number in the message header is the minimum version supported by the web-cache. Otherwise, if clear, this bit indicates that the protocol version number in the message header is the maximum version supported by the web-cache. This is used as part of the protocol version negotiation (see Section 3.4).

Bits 4 to 15:

Reserved, must be zero.

Assignment Data

The format of Assignment Data is specified by the setting of the Type bits within the Flags field, as follows:

Hash Assignment:

Hash Assignment Data Element (Section 6.6)

Mask Assignment:

Mask Assignment Data Element (Section 6.7)

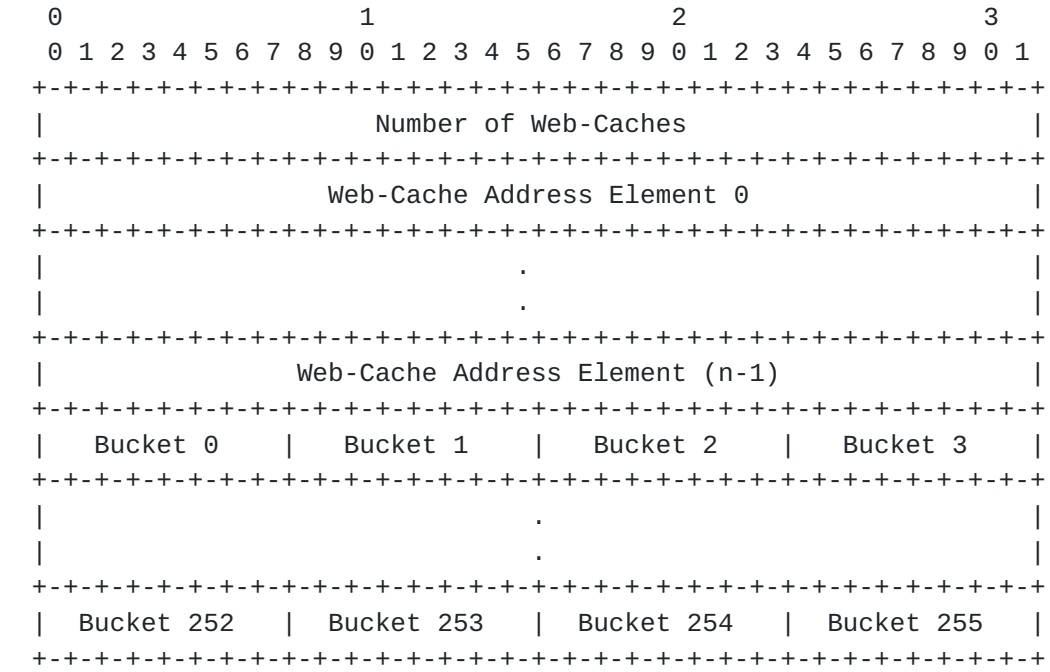
No Assignment:

The Assignment Data field is not present.

Extended Assignment:

Extended Assignment Data Element (Section 6.10)

6.5. Hash Buckets Assignment Element



Number of Web-Caches

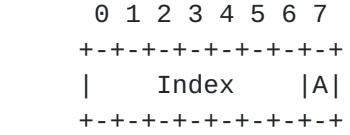
The number of useable web-caches (n) in the Service Group seen by all routers.

Web-Cache Address Element 0 -> Web-Cache Address Element (n-1)

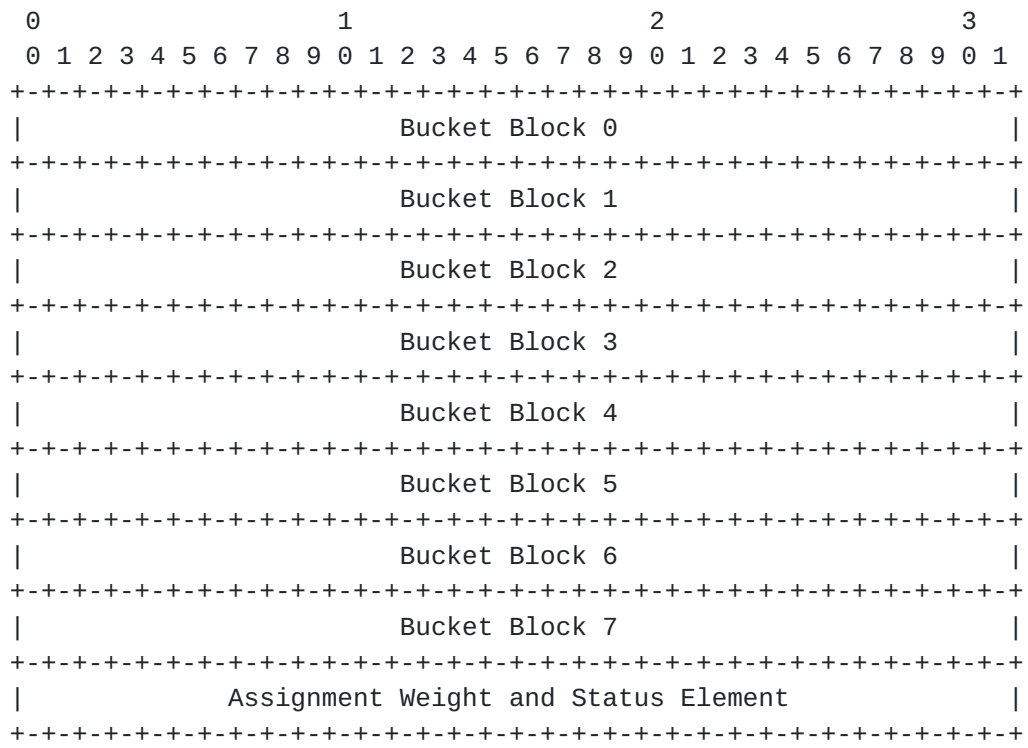
Elements indicating the IP addresses of the useable web-caches in the Service Group. The position of a web-cache in this list is the web-cache index. The first entry in the list has an index of 0. Each address element is defined in Section 4.7.

Bucket 0 -> Bucket 255

Contents of the Redirection Hash Table. The content of each bucket is a web-cache index value in the range 0 to 31. If set, the "A" flag indicates that alternative hashing should be used for this web-cache. The special value 0xFF indicates that no web-cache has been assigned to the bucket.



6.6. Hash Assignment Data Element



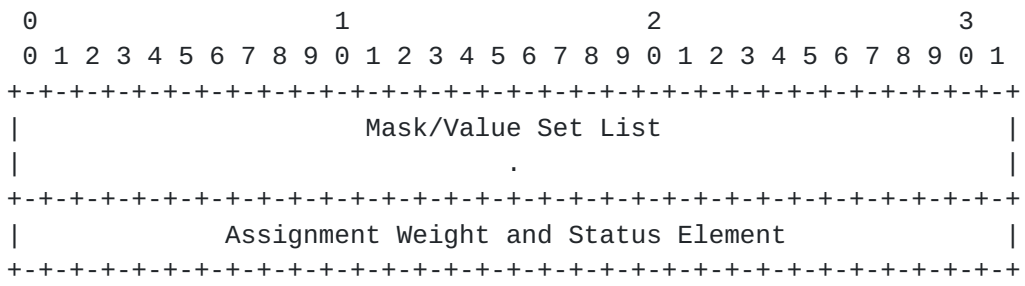
Bucket Block 0 -> Bucket Block 7

A 256-bit vector. A set bit indicates that the corresponding Redirection Hash Table bucket is assigned to this web-cache.

Assignment Weight and Status Element

This element may be used to indicate to the designated web-cache how new assignments should be made. The element is defined in Section 6.9.

6.7. Mask Assignment Data Element



Mask/Value Set List

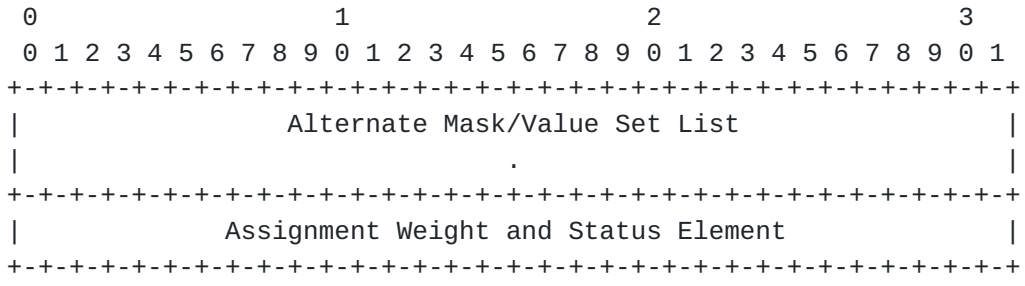
A list of mask/value sets. The list is defined in Section 6.13.

Assignment Weight and Status Element

This element may be used to indicate to the designated web-cache how new assignments should be made. The element is defined in Section 6.9.

6.8. Alternate Mask Assignment Data Element

This element provides a more compact representation of mask assignment data than the Mask Assignment Data Element. The Alternate Mask Assignment Data Element should be used in preference to the Mask Assignment Data Element whenever possible.



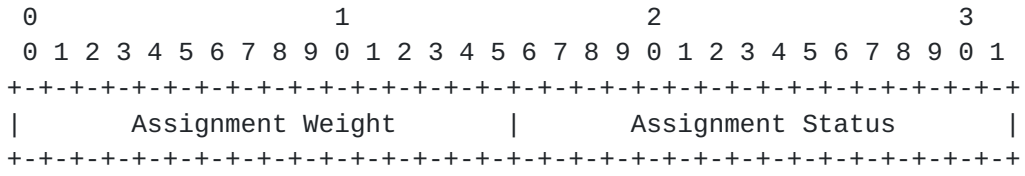
Alternate Mask/Value Set List

A list of alternate mask/value sets. The list is defined in Section 6.17.

Assignment Weight and Status Element

This element may be used to indicate to the designated web-cache how new assignments should be made. The element is defined in Section 6.9.

6.9. Assignment Weight and Status Element



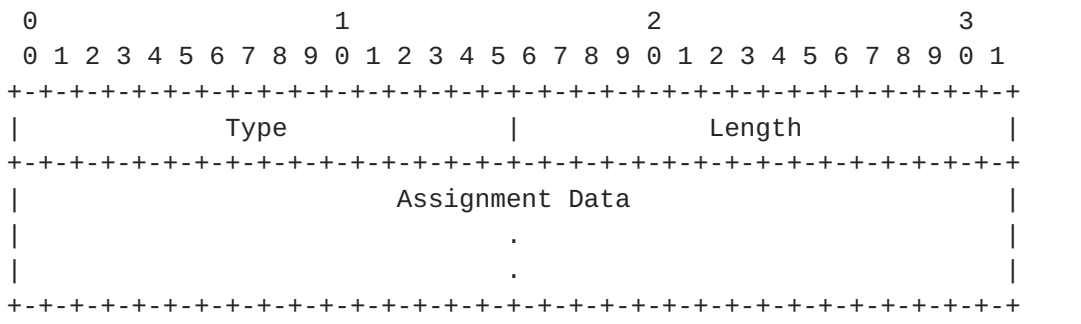
Assignment Weight

May be used to indicate to the designated web-cache how new assignments should be made. This information is generated by each web-cache to be associated with its identity information. It is received, stored and distributed by a router without modification.

Assignment Status

May be used to indicate to the designated web-cache how new assignments should be made. This information is generated by each web-cache to be associated with its identity information. It is received, stored and distributed by a router without modification.

6.10. Extended Assignment Data Element



Type

Indicates the format of Assignment Data. The currently defined values are:

0x00 - WCCP2_HASH_ASSIGNMENT
0x01 - WCCP2_MASK_ASSIGNMENT
0x02 - WCCP2_ALT_MASK_ASSIGNMENT
0x03 - WCCP2_ASSIGNMENT_WEIGHT_STATUS

Length

Length of the remainder of the element (Assignment Data).

Assignment Data

The format of Assignment Data is specified by the value of Type, as follows:

WCCP2_HASH_ASSIGNMENT:

Hash Assignment Data Element (Section 6.6)

WCCP2_MASK_ASSIGNMENT:

Mask Assignment Data Element (Section 6.7)

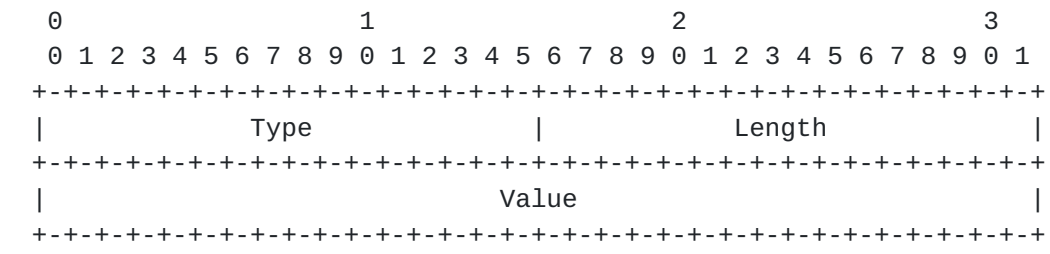
WCCP2_ALT_MASK_ASSIGNMENT:

Alternate Mask Assignment Data Element (Section 6.8)

WCCP2_ASSIGNMENT_WEIGHT_STATUS:

Assignment Weight and Status Element (Section 6.9)

6.11. Capability Element



Type

Currently defined types are:

0x01 - WCCP2_FORWARDING_METHOD (Section 6.11.1)
0x02 - WCCP2_ASSIGNMENT_METHOD (Section 6.11.2)
0x03 - WCCP2_PACKET_RETURN_METHOD (Section 6.11.3)
0x04 - WCCP2_TRANSMIT_T (Section 6.11.4)
0x05 - WCCP2_TIMER_SCALE (Section 6.11.5)

Routers and web-caches must ignore any Capability Element which has an unrecognised type.

Length

The length in octets of the following Capability Element Value.

Value

The format and length of the Value field is determined by the capability type. The following sub-sections describe the format of this field for each defined type.

6.11.1. Capability Type WCCP2_FORWARDING_METHOD

The Capability Element Value contains a 32-bit bitmask indicating the supported or selected forwarding methods. The currently defined values are:

0x00000001 - WCCP2_FORWARDING_METHOD_GRE
0x00000002 - WCCP2_FORWARDING_METHOD_L2

6.11.2. Capability Type WCCP2_ASSIGNMENT_METHOD

The Capability Element Value contains a 32-bit bitmask indicating the supported or selected assignment methods. The currently defined values are:

0x00000001 - WCCP2_ASSIGNMENT_METHOD_HASH
0x00000002 - WCCP2_ASSIGNEMNT_METHOD_MASK

6.11.3. Capability Type WCCP2_PACKET_RETURN_METHOD

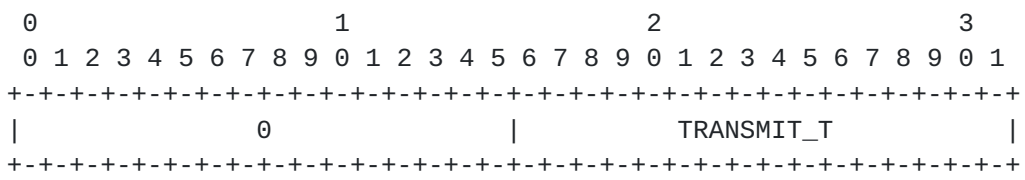
The Capability Element Value contains a 32-bit bitmask indicating the supported or selected packet return methods. The currently defined values are:

0x00000001 - WCCP2_PACKET_RETURN_METHOD_GRE
0x00000002 - WCCP2_PACKET_RETURN_METHOD_L2

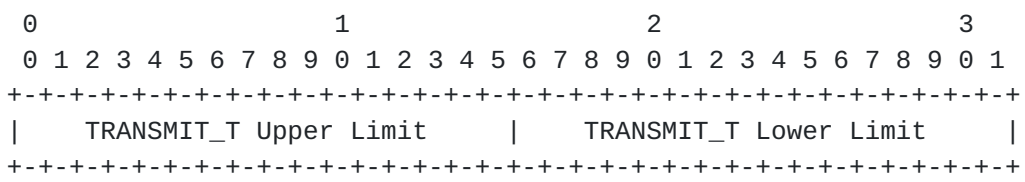
6.11.4. Capability Type WCCP2_TRANSMIT_T

The Capability Element Value contains two 16-bit values specifying the supported or selected TRANSMIT_T message interval in milliseconds. In a WCCP2_I_SEE_YOU message, a router can advertise either a range of permitted TRANSMIT_T values, or a single permitted TRANSMIT_T value. In a WCCP2_HERE_I_AM message, a web-cache can select only a single TRANSMIT_T value.

When a single selected value is to be specified, the first 16-bit value is zero and the second 16-bit value is the selected TRANSMIT_T message interval value:



When a supported range of permitted values is to be specified, the first 16-bit value contains the upper limit of the range and the second 16-bit value contains the lower limit of the range:



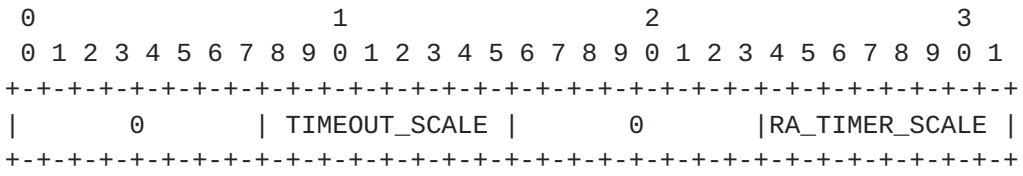
The default TRANSMIT_T value is 10000 (10 seconds) and applies when the WCCP2_TRANSMIT_T capability is not present. The range of supported values may be chosen by the implementation, but a minimum value of 500 and a maximum value of 60000 are suggested.

6.11.5. Capability Type WCCP2_TIMER_SCALE

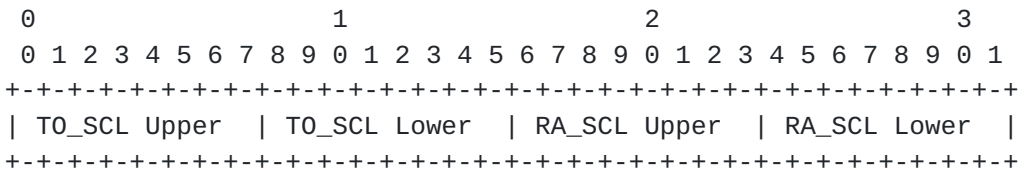
The Capability Element Value contains four 8-bit values specifying the supported or selected TIMEOUT_SCALE and RA_TIMER_SCALE values. In a WCCP2_I_SEE_YOU message, a router can advertise either a range of supported values for each parameter, or a single value for each parameter. In a WCCP2_HERE_I_AM message, a web-cache can select only a single value for each parameter.

The first and second 8-bit values are used to specify the TIMEOUT_SCALE parameter. The third and fourth 8-bit values are used to specify the RA_TIMER_SCALE parameter.

When a single selected value is to be specified for each parameter, the first 8-bit value is zero, the second 8-bit value is the selected TIMEOUT_SCALE value, the third 8-bit value is zero and the fourth 8-bit value is the selected RA_TIMER_SCALE value:



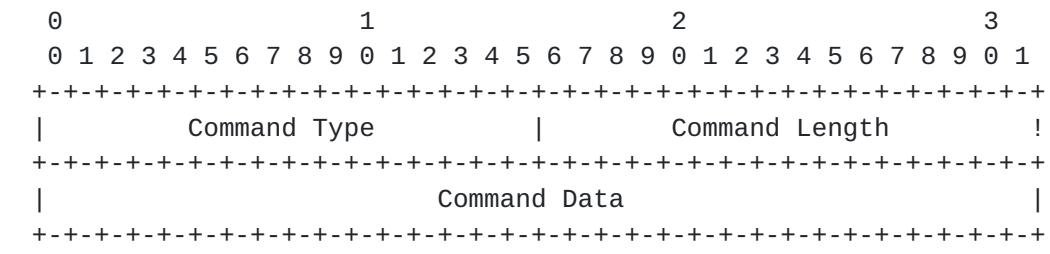
When a supported range of permitted values is to be specified for each parameter, the first 8-bit value contains the upper limit of the TIMEOUT_SCALE range, the second 8-bit value contains the lower limit of the TIMEOUT_SCALE range, the third 8-bit value contains the upper limit of the RA_TIMER_SCALE range and the fourth 8-bit value contains the lower limit of the TIMEOUT_SCALE range:



TO_SCL Upper = TIMEOUT_SCALE Upper Limit
TO_SCL Lower = TIMEOUT_SCALE Lower Limit
RA_SCL Upper = RA_TIMER_SCALE Upper Limit
RA_SCL Lower = RA_TIMER_SCALE Lower Limit

The default TIMEOUT_SCALE and RA_TIMER_SCALE values are both 1 and apply when the WCCP2_TIMER_SCALE capability is not present. The range of supported values for each of these parameters may be chosen by the implementation, but a minimum value of 1 and a maximum value of 5 are suggested in both cases. The value 0 must not be within the supported range of either parameter.

6.12. Command Element



Command Type

Currently defined command types are:

0x01 - WCCP2_COMMAND_TYPE_SHUTDOWN (Section 6.12.1)
0x02 - WCCP2_COMMAND_TYPE_SHUTDOWN_RESPONSE (Section 6.12.2)

Routers and web-caches must ignore any Command Element which has an unrecognised type.

Command Length

The length in octets of the following Command Data field.

Command Data

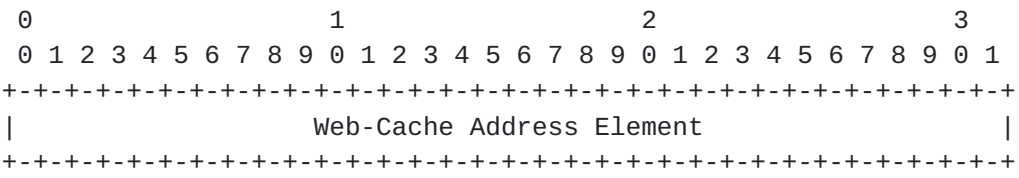
The format and length of the Command Data field is determined by the value of the Command Type field. The following sub-sections describe the format of this field for each defined type.

6.12.1. Command Type WCCP2_COMMAND_TYPE_SHUTDOWN

This command is used by a web-cache to indicate to the routers in a Service Group that it is shutting down and should no longer receive any redirected traffic.

The Command Data for the WCCP2_COMMAND_TYPE_SHUTDOWN command is a Web-cache IP address element, as defined in Section 4.7. The length of the field is 4 octets.

The format of the Command Data field is:



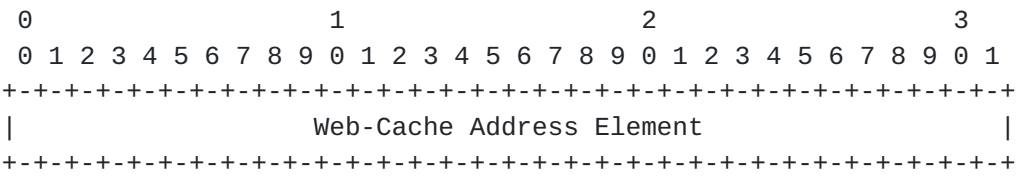
The address element value will be identical to that used in the Web-Cache Identity Element within the Web-Cache Identity Info Component.

6.12.2. Command Type WCCP2_COMMAND_TYPE_SHUTDOWN_RESPONSE

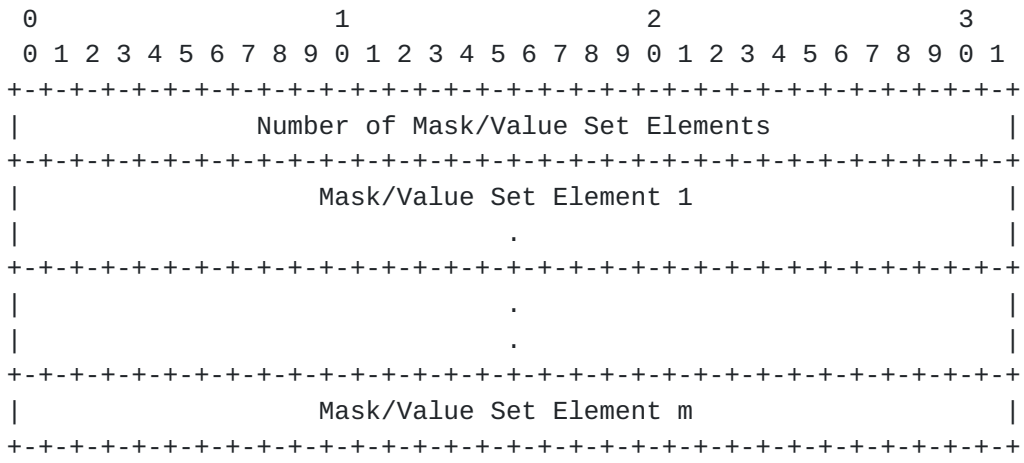
This command is used by a router to acknowledge receipt of a SHUTDOWN command received from the web-cache identified by the IP address element in the Command Data field.

The Command Data for the WCCP2_COMMAND_TYPE_SHUTDOWN_RESPONSE command is a Web-cache IP address element, as defined in Section 4.7. The length of the field is 4 octets.

The format of the Command Data field is:



6.13. Mask/Value Set List



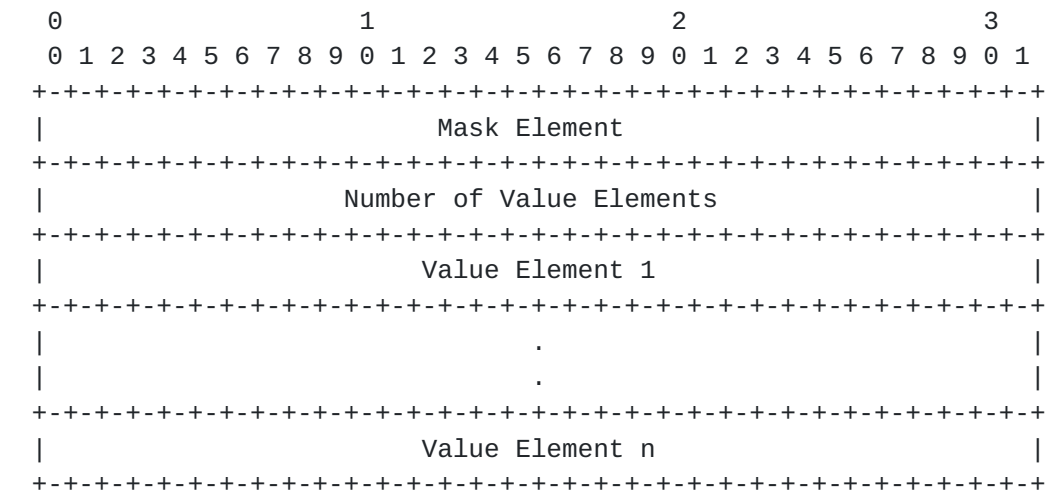
Number of Mask/Value Set Elements

The number of Mask/Value Set Elements (m) in the following list.

Mask/Value Set Element 1 -> Mask/Value Set Element m

A list of the Mask/Value Set Elements. Each element is defined in Section 6.14.

6.14. Mask/Value Set Element



Mask Element

The Mask Element for this set. The element is defined in Section 6.15.

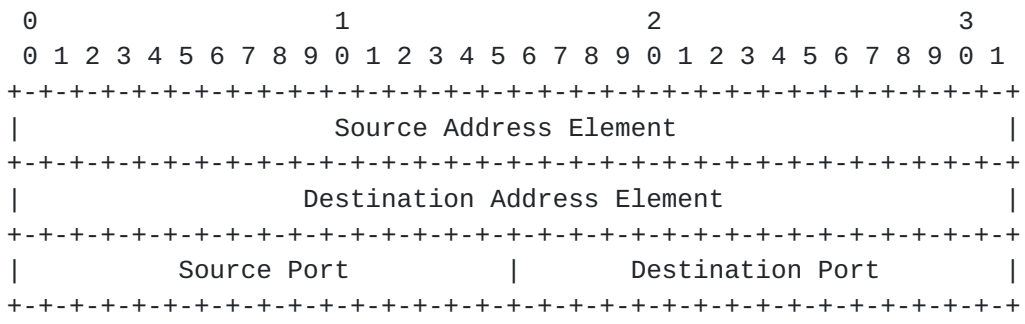
Number of Value Elements

The number of Value Elements (n) in this set.

Value Element 1 -> Value Element n

The Value Elements for this set. Each element is defined in Section 6.16.

6.15. Mask Element



Source Address Element

Indicates the mask to be applied to the source IP address of the packet. A value of zero means "Don't care". The element is defined in Section 4.7.

Destination Address Element

Indicates the mask to be applied to the destination IP address of the packet. A value of zero means "Don't care". The element is defined in Section 4.7.

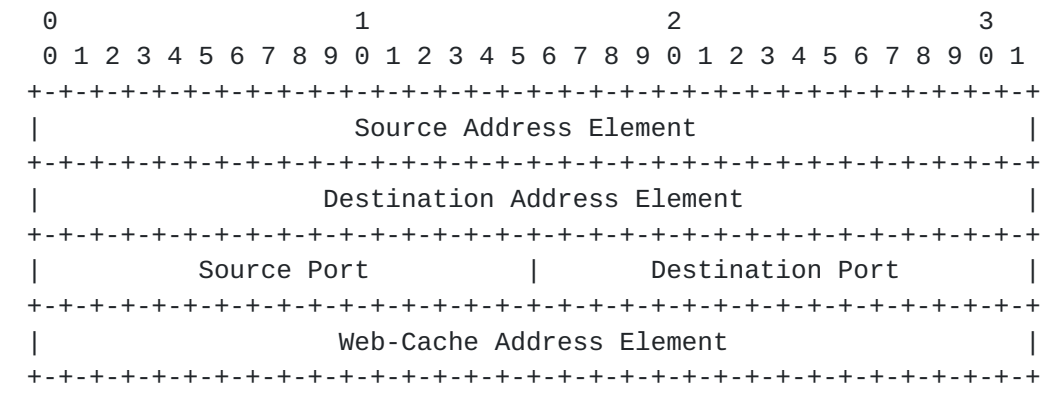
Source Port

The 16-bit mask to be applied to the TCP/UDP source port field of the packet. A value of zero means "Don't care".

Destination Port

The 16-bit mask to be applied to the TCP/UDP destination port field of the packet. A value of zero means "Don't care".

6.16. Value Element



Source Address Element

Indicates the value to match against the source IP address of the packet after masking. The element is defined in Section 4.7.

Destination Address Element

Indicates the value to match against the destination IP address of the packet after masking. The element is defined in Section 4.7.

Source Port

The value to match against the TCP/UDP source port number of the packet after masking.

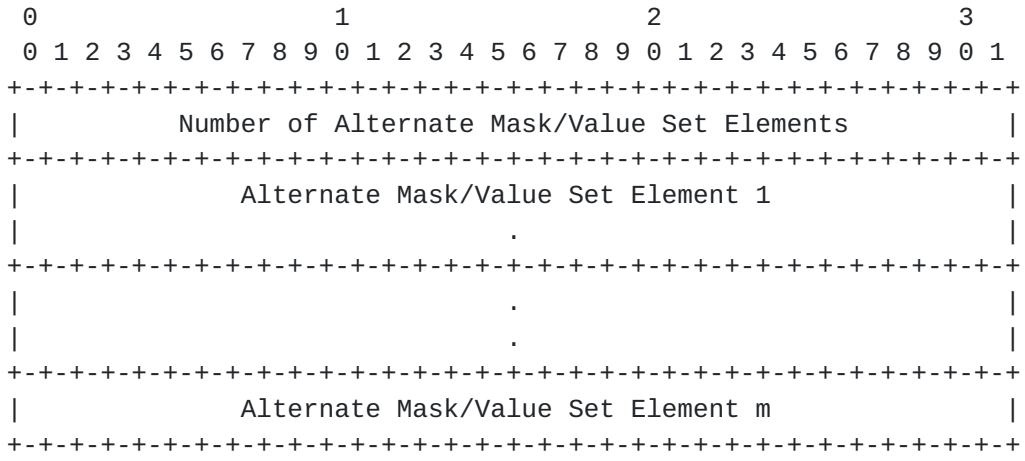
Destination Port

The value to match against the TCP/UDP destination port number of the packet after masking.

Web-Cache Address Element

Indicates the identifying IP address of the web-cache to which packets matching this Value Element should be sent. The address element is defined in Section 4.7.

6.17. Alternate Mask/Value Set List



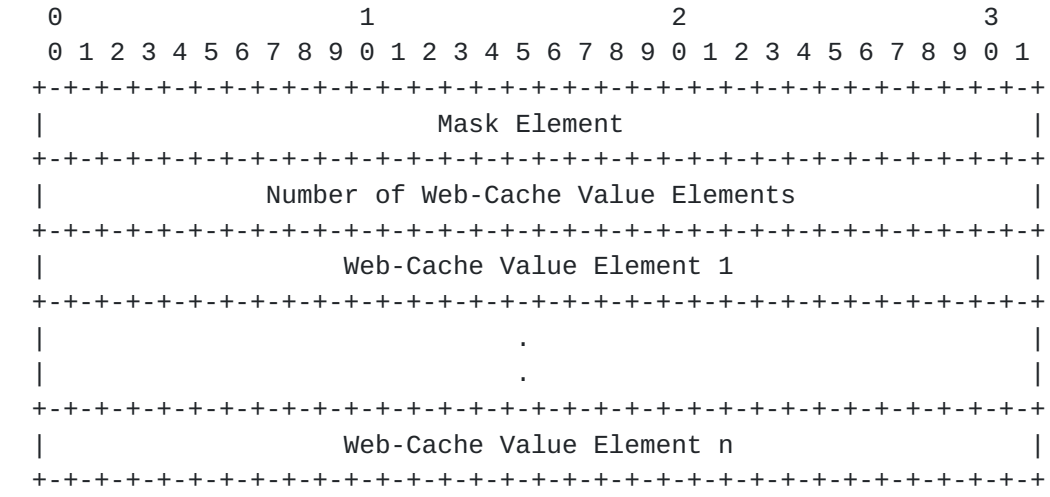
Number of Alternate Mask/Value Set Elements

The number of Alternate Mask/Value Set Elements (m) in the following list.

Alternate Mask/Value Set Element 1 -> Alternate Mask/Value Set Element m

A list of Alternate Mask/Value Set Elements. Each element is defined in Section 6.18.

6.18. Alternate Mask/Value Set Element



Mask Element

The Mask Element for this set. The element is defined in Section 6.15.

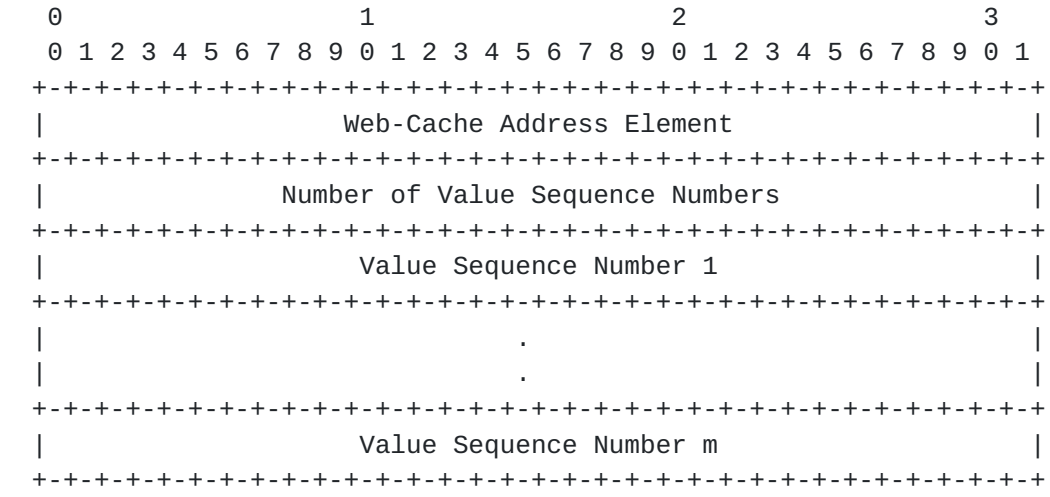
Number of Web-Cache Value Elements

The number of Web-cache Value Elements in this set.

Web-Cache Value Element 1 -> Web-Cache Value Element n

The Web-cache Value Elements for this set. Each element is defined in Section 6.19.

6.19. Web-Cache Value Element



Web-Cache Address Element

Indicates the identifying IP address of the web-cache to which packets matching this list of value sequence numbers should be sent. The address element is defined in Section 4.7.

Number of Value Sequence Numbers

The number of Value Sequence Numbers (m) in this element.

Value Sequence Number 1 -> Value Sequence Number m

An index (starting from 0) into an imaginary table that contains an entry for each possible value that could be matched against the result of applying the mask to the fields of the packet header. The size of the imaginary table is determined by the total number of bits set in the mask. For n bits set in the mask, the imaginary table contains 2^n (2 raised to the power n) entries. The minimum permitted index value is 0 and the maximum permitted index value is (2^n)-1.

7. Interpreting Alternate Mask/value Set Elements

As defined in Section 6.15, each mask consists of four elements:

1. Source address mask (SAM)
2. Destination address mask (DAM)
3. Source port mask (SPM)
4. Destination port mask (DPM)

Each bit that is set in any of the four mask elements maps uniquely to an individual bit within the Value Sequence Number (VSN). With 32 bits available in the VSN, there can be up to 32 bits set in the mask across the four elements.

The order of the mask elements listed above is the order of significance, with the SAM being the most significant element (MSE) and the DPM being the least significant element (LSE).

Bits within the VSN are mapped in order from the least significant bit (LSB, bit 0) to the most significant bit (MSB, bit 31). Mask elements are processed in order from the LSE to the MSE. Within each mask element, octets are processed from the least significant octet to the most significant octet, and within each octet bits are processed from the LSB (bit 0) to the MSB (bit 7).

For example, consider the following IPv4 mask:

Source Address Mask	Dest Address Mask	Source Port Mask	Dest Port Mask
0x00000100	0x00000003	0x0000	0x0001

When mapping bits in the mask above to bits in the VSN, the values shown above are processed from right to left as follows.

The least significant element is the DPM. Within that element, bit 0 is set in the least significant octet, therefore this is mapped to bit 0 in the VSN. No other bits are set within the DPM, so processing moves on to the SPM.

No bits are set in the SPM so processing moves on to the DAM.

In the least significant octet of the DAM, bit 0 is set therefore this is mapped to the next available bit in the VSN, bit 1. The next bit set in the DAM is bit 1 of the least significant octet, so it maps to the next available bit in the VSN, bit 2. No other bits are set within the DAM, so processing moves on to the SAM.

In the least significant octet of the SAM, no bits are set, so processing moves on to the next significant octet within the SAM. In this octet, bit 0 is set therefore this is mapped to the next available bit in the VSN, bit 3.

Therefore, the above mask results in the following mapping (mask octets are counted from least significant to most significant):

VSN bit 0 --> DPM octet 0, bit 0
VSN bit 1 --> DAM octet 0, bit 0
VSN bit 2 --> DAM octet 0, bit 1
VSN bit 3 --> SAM octet 1, bit 0

Using the mapping shown above, the following table can be constructed. It shows the values that correspond to each valid VSN:

Value Sequence Number	Source Address Value	Dest Address Value	Source Port Value	Dest Port Value
0	0x00000000	0x00000000	0x0000	0x0000
1	0x00000000	0x00000000	0x0000	0x0001
2	0x00000000	0x00000001	0x0000	0x0000
3	0x00000000	0x00000001	0x0000	0x0001
4	0x00000000	0x00000002	0x0000	0x0000
5	0x00000000	0x00000002	0x0000	0x0001
6	0x00000000	0x00000003	0x0000	0x0000
7	0x00000000	0x00000003	0x0000	0x0001
8	0x00000100	0x00000000	0x0000	0x0000
9	0x00000100	0x00000000	0x0000	0x0001
10	0x00000100	0x00000001	0x0000	0x0000
11	0x00000100	0x00000001	0x0000	0x0001
12	0x00000100	0x00000002	0x0000	0x0000
13	0x00000100	0x00000002	0x0000	0x0001
14	0x00000100	0x00000003	0x0000	0x0000
15	0x00000100	0x00000003	0x0000	0x0001

The table above is equivalent to a list of all possible values which can be obtained by applying the mask to any input data, arranged into a specific sequential order. For the given mask, each VSN is effectively an index into this table. However, to convert between a VSN and its equivalent value, a table lookup is not required as the preceding bit mapping achieves the same result.

In an Alternate Mask/Value Set Element, each web-cache is represented by a Web-Cache Value Element. For each web-cache there is a list of VSNs within the Web-Cache Value Element to show which values have been assigned to the web-cache.

For example, in an Alternate Mask/Value Set Element listing three web-caches, each may have a list of VSNs as follows:

web-cache 1, VSNs: 0, 3, 6, 9, 12, 15
web-cache 2, VSNs: 1, 4, 7, 10, 13
web-cache 3, VSNs: 2, 5, 8, 11, 14

This is equivalent to the following values in a Mask/Value Set Element:

Source Address Value	Dest Address Value	Source Port Value	Dest Port Value	Target Web-cache
0x00000000	0x00000000	0x0000	0x0000	1
0x00000000	0x00000000	0x0000	0x0001	2
0x00000000	0x00000001	0x0000	0x0000	3
0x00000000	0x00000001	0x0000	0x0001	1
0x00000000	0x00000002	0x0000	0x0000	2
0x00000000	0x00000002	0x0000	0x0001	3
0x00000000	0x00000003	0x0000	0x0000	1
0x00000000	0x00000003	0x0000	0x0001	2
0x00000100	0x00000000	0x0000	0x0000	3
0x00000100	0x00000000	0x0000	0x0001	1
0x00000100	0x00000001	0x0000	0x0000	2
0x00000100	0x00000001	0x0000	0x0001	3
0x00000100	0x00000002	0x0000	0x0000	1
0x00000100	0x00000002	0x0000	0x0001	2
0x00000100	0x00000003	0x0000	0x0000	3
0x00000100	0x00000003	0x0000	0x0001	1

In the example above, all valid VSNs are used but this is not a requirement, each VSN does not need to be assigned to a web-cache. However, it is a requirement that each VSN is listed for no more than one web-cache.

Generally, as demonstrated above, Alternate Mask/Value Set Lists can be used to represent the same information as Mask/Value Set Lists, but in a more compact form. Therefore, when constructing a WCCP message in which protocol version 2.01 is used, Alternate Mask/Value Set Lists should be used in preference to Mask/Value Set Lists to achieve a smaller message size.

8. Security Considerations

WCCP V2 provides a mechanism for message authentication. It is described in Section 3.7 of this document. The authentication mechanism relies on a password known to all routers and web-caches in a Service Group. The password is part of the Service Group configuration and is used to compute message checksums which can be verified by other members of the group. Should the password become known to a host attempting to disrupt the operation of a Service Group it would be possible for that host to spoof WCCP messages and appear as either a router or web-cache in the Service Group.

To pose as a router in a Service Group a host would advertise its presence to the members of the group in I_SEE_YOU messages. If accepted as part of the Service Group the host would receive the configuration for the group in a HERE_I_AM message from the designated web-cache. This situation would not pose any threat to the operation of the Service Group because the host would not be performing any packet redirection and all packets would flow normally.

To pose as a web-cache within a Service Group a host would advertise its presence in HERE_I_AM messages. Acceptance of the host as part of the Service Group would be decided by the designated web-cache and may be subject to additional security checks not specified by WCCP. The host may attempt to become the designated web-cache to avoid these checks, but acceptance of a host as the designated web-cache may also be subject to additional security checks. Should the host become part of the Service Group it would be assigned a proportion of the traffic redirected by the routers in the Service Group. Assuming that the host drops any redirected packets, the net effect to clients would be the loss of a proportion of the traffic flowing through the Service Group routers.

9. IANA Considerations

This document has no actions for IANA.

10. Acknowledgements

The author would like to thank Martin Cieslak, Richard Edmonstone, Mark Gillott and Khalid Rafiq for their assistance in reviewing this document or earlier versions.

11. Normative References

[IANA-AF] Internet Assigned Numbers Authority, "Address Family Numbers", <http://www.iana.org/assignments/address-family-numbers>.

[RFC1321] Rivest, R., "The MD5 Message-Digest Algorithm", RFC 1321, April 1992.

[RFC1701] Hanks, S., Li, T., Farinacci, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 1701, October 1994.

Author's Address

Douglas J. McLaggan
Cisco Systems
96 Commercial Street
Edinburgh, EH6 6LX
United Kingdom

Email: djmclaggan@gmail.com