



# Behavioral Change for Buffer Recarving

---

**Part Number:** OL-18534-01

**First Published:** December 8, 2008

**Table 1 Feature History**

Cisco IOS Release	Modification
12.0(32.03)S03, 12.0(32)S08, 12.0(32)SY04	Introduced a new command-line interface (CLI) command “[no] bfrp carve-level [0   1]” to enable or disable packet buffer recarve optimization. The packet buffer recarving occurs when an MTU (maximum transmission unit) is changed on the router.
12.0(32)SY07, 12.0(32)S06c	The bug id caveat CSCso80492 introduced a behavioral change to the carve-level 1 mechanism. After changing the largest MTU configured on a slot, sometimes an extra buffer queue is added in the ToFab of other line cards. This additional of an extra buffer queue that is not in the global buffer distribution (dsn) is used to process packets that enter its range.
12.0(32)SY07, 12.0(32)S06c, and 12.0(33)S02	The bug id caveat CSCsq94862 introduced a behavioral change to the CLI command “no bfrp carve-level”.

## Contents

- [Purpose of This Document](#)
- [Overview](#)
- [Buffer Carving Algorithm](#)
- [Controlling Buffer Recarving](#)
- [Behavioral Change to Carve-Level 1](#)
- [Behavioral Change to Command “no bfrp carve-level”](#)



---

**Americas Headquarters:**  
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

# Purpose of This Document

This document explains the basics of the buffer recarving functionality, buffer carves that are triggered upon changing MTUs, the behavioral change to carve-level 1 functionality, the addition of the new command **bfrp carve-level** to enable or disable packet buffer recarve optimization, and behavioral change to the command “**no bfrp carve-level**”.

## Overview

Each line card (LC) on a Cisco 12000 series Internet router has two types of memory:

- Route or processor memory (dynamic RAM [DRAM]): This memory enables mainly the onboard processor to run Cisco IOS software and store network routing tables (Forwarding Information Base [FIB] adjacency).
- Packet memory (synchronous dynamic RAM [SDRAM]): Line card packet memory temporarily stores data packets awaiting switching decisions by the line card processor.

This document discusses mainly the buffer carving algorithm, the addition of a new CLI command to control buffer optimization, a behavioral change to carve-level 1 functionality and a behavioral change to the command **no bfrp carve-level**. Buffer carving mechanism is related with the packet memory (SDRAM), which is divided into two banks: ToFab and FrFab (towards the Fabric and from the Fabric). The ToFab memory is used for packets that come in one of the interfaces on the LC and make their way to the fabric, whereas the FrFab memory is used for packets that are going out of an interface on the LC from the fabric.

The ToFab and FrFab queues are the most important concepts to understand in order to efficiently troubleshoot ignored packets in the Cisco 12000 Series Internet Router.



---

**Note**

ToFab (toward the Fabric) and "Rx" (received by the router) are two different names for the same thing, as are "FrFab" (from the Fabric) and "Tx" (transmitted by the router). For example, the ToFab Buffer Management ASIC (BMA) is also referred to as the RxBMA. This document uses the ToFab or FrFab convention, but you may see the Rx or TX nomenclature used elsewhere.

---

## Buffer Carving Algorithm

Cisco IOS software implements a buffer-carving algorithm that divides SDRAM into various buffer sizes. The Cisco 12000 series Internet router connects to several different media types, with different MTUs. A mechanism was needed to make the most efficient use of packet memory buffers. The buffer carving algorithm was devised after studying the makeup of Internet traffic and considering the various media types that the Cisco 12000 series Internet router supports.

## Understanding Buffer Carving Mechanism

The Route Processor (RP) runs the buffer carve code to determine the buffer specifications to be sent down to line cards. Line cards apply these specifications on their packet buffers. Route Processor buffer carve code can be invoked under the following conditions:

- Router power-on
- MTU changes on an LC interface, done from the RP
- Online insertion and removal (OIR) event; for example, insertion or removal of a line card

## Controlling Buffer Recarving

Each line card on Cisco 12000 series Internet router has two sets of packet memories. One of the packet memories is on the ingress side (tofab) and the other is on the egress side (frfab).

The Route Processor calculates buffer specifications and sends it down to the line cards on their power-on. The line cards blindly carve their packet memories based on these specifications. Changes in the buffer specification can happen when the MTU distribution changes on the router, triggering recarving of buffers on line cards (LC). Currently by default, buffer recarving occurs in any of the following cases:

- Case 1: Change that results from increasing or decreasing the maximum MTU configured on the router.
- Case 2: Change that results from adding or removing new buffer pool within the maximum slot MTU, with no change in the maximum MTU on the router.
- Case 3: Change that results from increasing or decreasing the existing pool buffer length while the maximum MTU on the router remains the same.
- Case 4: Change that results from increasing or decreasing the maximum slot MTU, which may change the maximum MTU on the router.

## Network Traffic Disruption

Recarves on line card result in disruption of the traffic flow; the worst-case recarve will result in all line card Tofab and Fr Fab packet memories getting recarved, resulting in a router-wide traffic drop.

Recarving in [case 1](#) is acceptable to ensure the biggest packet flowing through the router does not get dropped. Similarly, [case 4](#) is acceptable because it ensures that the biggest packets on the ingress side do not get dropped. However, [case 2](#) and [case 3](#) can be disabled using the `bfrp carve-level` command.

## bfrp carve-level

To ensure control over the frequency of buffer recarving, use the **bfrp carve-level** command. Whenever buffer recarving occurs, packets sent to that line card are dropped. Use the command **bfrp carve-level 0** if you want to disable buffer recarving completely after recarving once to default buffer specifications. Use the command **bfrp carve-level 1** if you want to perform buffer recarving only if the change in the MTU is changing the maximum slot MTU configured on any slot. Use the **no bfrp carve-level** command to restore the default carving mechanism.

**bfrp carve-level [0 | 1]**

**no bfrp carve-level**

Syntax Description	0	Applies default buffer distribution and disables buffer recarves completely.
	1	Allows recarving only if the change in the MTU is changing the maximum slot MTU configured on any slot.

**Command Modes** Global configuration

**Defaults** By default, buffer recarving is triggered if any of the cases discussed in the section [“Controlling Buffer Recarving”](#) are met.

Command History	Release	Modification
	12.0(32.03)S03, 12.0(32)S08, 12.0(32)SY04	This command was introduced.

**Examples** The following example shows how to recarve the default buffer distribution and disable further recarving for any MTU change. The maximum MTU configured on router is 6000:

```
Router> enable
Router# config t
Router(config)# bfrp carve-level 0
```

The following example shows that recarving is invoked only if the MTU change is changing the maximum slot MTU configured on any slot:

```
Router(config)# bfrp carve-level 1
```

The following example shows how to restore the default buffer recarving behavior:

```
Router(config)# no bfrp carve-level
```

**Usage Guidelines** Use the command **bfrp carve-level 0** to accommodate all default buffer sizes on the router with a maximum buffer size of 9248 and then prevent further recarves on any MTU change.

Use the command **bfrp carve-level 1** to selectively disable buffer recarve, by allowing recarves only for maximum slot MTU changes.

Use the command **no bfrp carve-level** to restore to default carving mechanism.

**Note**

---

At any given time, only one of the two buffer levels (carve level 0 or carve level 1) can be activated.

---

# Behavioral Change to Carve-Level 1

This section explains the behavioral change to carve-level 1 functionality. It contains the following topics:

- [Old Behavior](#)
- [New Behavior](#)

## Old Behavior

The ToFab and FrFab of all line cards match the queues in the global buffer distribution (dsn).

## New Behavior

At carve level 1, fix ensures change in largest MTU configured on slot, recarves both ToFab and FrFab buffers on the same line card and only FrFab buffers on other card, keeping the ToFab buffers of other card unchanged.

However, in some rare cases though global distribution changes due to MTU change, same will not be reflected in the ToFab of other line card due above restriction. So, there could be some extra buffer queues which are not in global buffer distribution.

Buffers are allocated dynamically for incoming packets. If incoming packet matches the criteria for extra queue found in ToFab direction, it will allocate buffer from that queue. So buffers in extra queue(in ToFab direction) do not go waste.

The following example shows the logs of caveat at carve level-1. The command output in the following example shows that an extra buffer queue is added in ToFab for slot 2. The example includes the following hardware and software configurations:

- A router with three line cards available at slots 0, 2, and 3, respectively.
- The default MTUs configured on slots 0, 2, and 3 are 4470, 1500, and 4470, respectively.
- The initial maximum slot MTUs configured on slots 0, 2, and 3 are 4470, 9072, and 1500, respectively. (These steps followed to produce the following output).

- 
- Step 1** At slot 3, change the maximum slot MTU to 1800 from the default value of 1500.
- Step 2** At slot 2, change the maximum slot MTU from 9072 to the default value of 4470.
- Step 3** At slot 3, change the maximum slot MTU from 1800 back to the default value of 1500.
- 

```
Router# show controllers queues
Global MTU information:
  Max MTU data size = 4592
  MTUs considered in SDRAM buffer carving:
    Index  MTU Data Size
    0      1568
    1      4592

Input SDRAM buffer distribution = default :
Index  Data Size      Min Data Size  Percent Buffers  Fixed?
  0      80              1              50              yes
  1     608              81              30              yes
```

2	1616	1488	15	yes
3	4592	3671	3	no
4	9296	7431	2	no

Resulting global SDRAM buffer distribution (includes MTUs):

Index	Data Size	Min Data Size	Percent Buffers	Fixed?
0	80	1	50	yes
1	608	81	30	yes
2	1616	1488	15	yes
3	4592	3671	3	no
4	9296	7431	2	no

Contd...

FrFab SDRAM buffer carve-data for slot 0:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	51
1	608	31
2	1616	15
3	4592	3

ToFab SDRAM buffer carve-data for slot 0:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	51
1	608	31
2	1616	15
3	4592	3

FrFab SDRAM buffer carve-data for slot 2:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	51
1	608	31
2	1616	15
3	4592	3

ToFab SDRAM buffer carve-data for slot 2:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	50
1	608	31
2	1616	15
3	1952	1 <<<< Mismatch with global dsn
4	4592	3

FrFab SDRAM buffer carve-data for slot 3:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	51
1	608	31
2	1616	15
3	4592	3

ToFab SDRAM buffer carve-data for slot 3:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	53
1	608	32
2	1616	16

# Behavioral Change to Command “no bfrp carve-level”

This section explains the behavioral change to the command “**no bfrp carve-level**”. It contains the following topics:

- [Old Behavior](#)
- [New Behavior](#)

## Old Behavior

Traversing back from carve level 0 or 1 to the default carve using the command “**no bfrp carve-level**”, restores to the default carving mechanism without performing any carving. Only future MTU changes trigger recarve of both ToFab and From Fab according to their [default carve mechanism](#).

## New Behavior

Executing the command “**no bfrp carve-level**”, triggers a carving on both ToFab and FromFab direction according to the default carving mechanism. The feature change is that if you traverse back from carve-level 0 or 1 to default carve mechanism, using the command “**no bfrp carve-level**”, recarving takes place. Below is the output showing the new behavior. It triggers router wide recarving, suggested to use on a maintenance window (MW).

- 
- Step 1** The following example shows how to view the current carve-level 0 or 1 configured on the router. Execute the following command from Global configuration mode.

```
GSR-12012(config)# do show running-config | i carve
bfrp carve-level 0
```

- Step 2** The following example shows how to traverse back from carve-level 0 or 1 to the default carve. The command output message indicates that there will be recarving of buffers.

```
GSR-12012(config)# no bfrp carve-level
GSR-12012(config)#
05:05:42: %RP-6-CARVE_BUF: Carve Level change may recarve the buffers << indicative
message
GSR-12012(config)#end
GSR-12012(config)#end
GSR-12012#
05:08:51: %SYS-5-CONFIG_I: Configured from console by console
```

- Step 3** The following example shows how to view the current global distribution list and the carves configured at various slots on the router. Execute the following command from Exec mode.

```
GSR-12012# show controllers queues <<< buffers are recarved according to default carve
mechanism
Global MTU information:
  Max MTU data size = 4592
  MTUs considered in SDRAM buffer carving:
    Index  MTU Data Size
    0      4592
Input SDRAM buffer distribution = default :
  Index  Data Size  Min Data Size  Percent Buffers  Fixed?
  0      80        1              50
yes
```



	1	608	81	30
yes				
	2	1616	1488	15
yes				
	3	4592	3671	3
no				
	4	9296	7431	2
no				

Contd...

Resulting global SDRAM buffer distribution (includes MTUs):

	Index	Data Size	Min Data Size	Percent Buffers	Fixed?
	0	80		1	50
yes					
	1	608	81	30	yes
	2	1616		1488	15
yes					
	3	4592		3671	3
no					
	4	9296		7431	2
no					

FrFab SDRAM buffer carve-data for slot 8:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	51
1	608	31
2	1616	15
3	4592	3

ToFab SDRAM buffer carve-data for slot 8:

valid element = yes, percent buffers unused = 0

Buffer distribution:

Index	Data Size	Percent Buffers
0	80	51
1	608	31
2	1616	15
3	4592	3

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, Cisco TelePresence, Cisco WebEx, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn and Cisco Store are service marks; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study, IronPort, the IronPort logo, LightStream, Linksys, MediaTone, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0809R)

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

© 2008 Cisco Systems, Inc. All rights reserved.