Understanding Queuing With Hierarchical Queueing Framework (HQF)

Last updated: June 2012

Several modifications to the current buffering capabilities have been made over the last few releases of IOS software. This has raised questions on the interworking of buffering/queueing in Hierarchical Queueing Framework (HQF) based systems. This document will discuss the current structure of buffering, configuration options, and detail the interworking of the necessary configurations with respect to the ISR platform.

It should be noted that I/O memory size will vary per box, which is tied to the overall buffering capability. This document will cover the buffering process and will use cli for verification purposes.

The Process of Buffering

Cisco routers store packets in a few different locations depending on the congestion level:

1. Tx-Ring—hw queue or transmit ring
2. Interface buffers—public or private pools
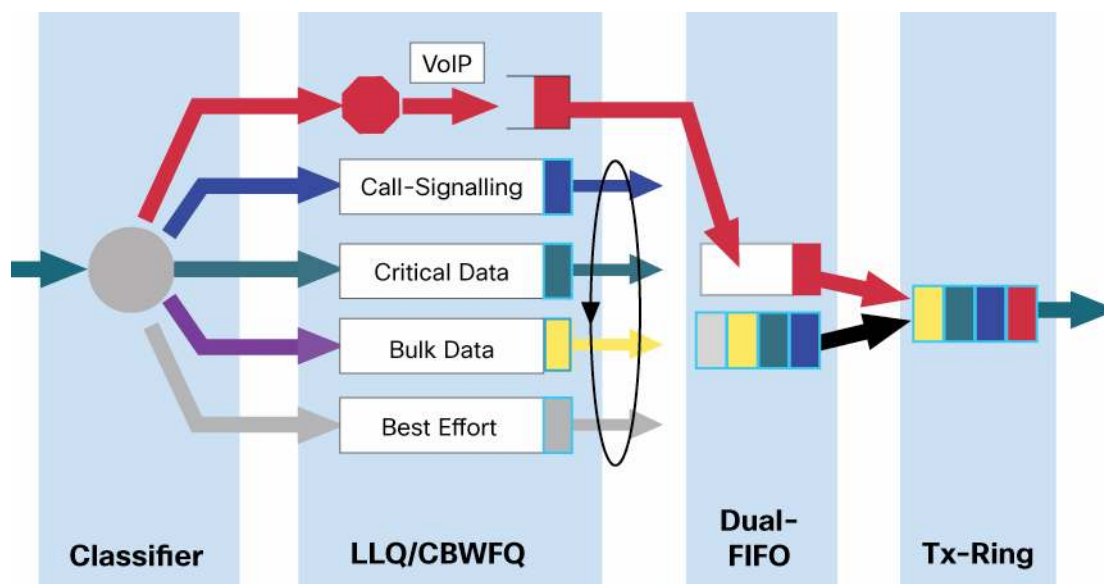
As a packet enters a router, the packet is stored inside of the ingress buffer waiting to be processed.

The packet when sent to the egress for processing is still accounted for on the ingress interface. If the packet has exceeded the 1k buffers on the physical port, memory will be taken from the public pool.

Overview of System Buffers

The public pools of the mid-range routing series offers two possible locations to store packets: the Fast Switching buffers (F/S) and the normal pool buffers (depending on the IOS version, these buffers are used as the backup pool when the private interface pool is filled with packets). These buffers vary in size and access speed.

You can verify the buffers with the <show buffers> command:

```
Public particle pools:
F/S buffers, 1664 bytes (total 1536, permanent 1536):
     1024 in free list (256 min, 2048 max allowed)
     512 hits, 0 misses, 0 trims, 0 created
     0 failures (0 no memory)
     512 max cache size, 0 in cache
     0 hits in cache, 0 misses in cache
Normal buffers, 1676 bytes (total 3840, permanent 3840):
     3840 in free list (128 min, 4096 max allowed)
     0 hits, 0 misses, 0 trims, 0 created
     0 failures (0 no memory)
```
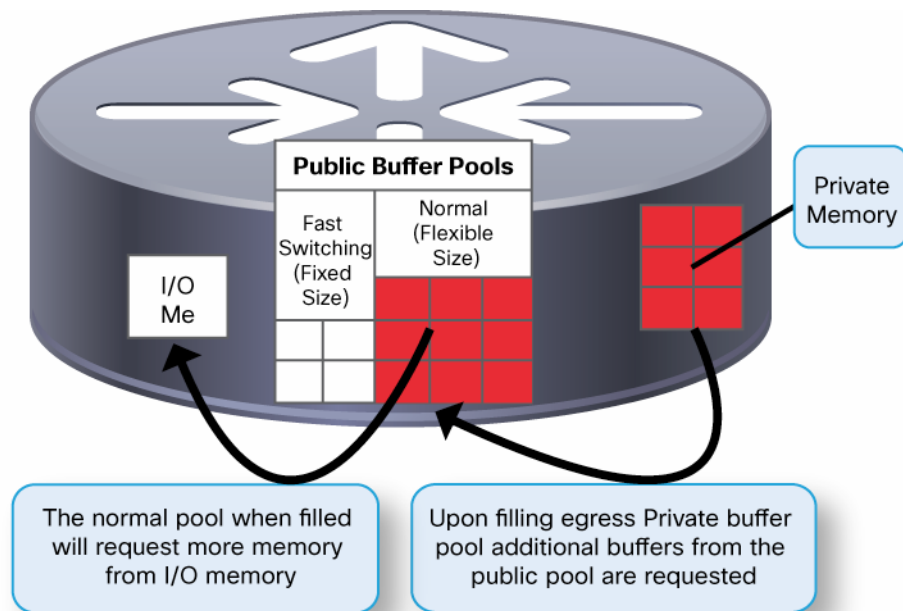
The public buffers (public particle pools) are used when the depth of the private buffers is reached.

The F/S buffer can provide memory for 2048 packets (512 packets in the cache and 1536 packets in the permanent memory). The soft limit is governed by the configured amount of fast switching buffers <buffers fastswitching permanent x>.

The normal buffers provide a maximum packet memory of 4096 packets (256 packets in the cache and 3840 packets in the permanent memory).

The ISR G1 and G2 platforms differ when using public buffer pools. The ISR G1 always uses the normal buffer pool. In IOS versions up to IOS 15.2T, the ISR G2 uses the Fast Switching buffer (F/S), beyond this release the normal pool buffers are used (CSCtq27141, CSCtw65356). This modification provides a larger backup pool that is useable for complex Quality of Service (QoS) configurations.

The public buffers are partially in cache, ready for use after the physical port is oversubscribed. At this point, if the router has filled up both the private and the cache public buffers, the router will request more public buffers from I/O memory. The upper value of physical I/O memory is governed by the amount of RAM on the platform.

The normal pool when filled will request more memory from I/O memory

Upon filling egress Private buffer pool additional buffers from the public pool are requested

HQF Architectural Integration

HQF (Hierarchical Queueing Framework) makes use of the physical buffering. The hold-queue parameter per interface governs the amount of buffer that the physical interface can use. The per-class queue-limit configuration limits the per-class buffering threshold. These values must be in alignment with the physical hardware limits, otherwise packet loss will occur.

Note: The aggregate queue-limit of the configured CBWFQ classes must be equal or smaller than the configured <hold-queue *packets* out> configured on the assigned interface. If the hold-queue limit is smaller than the configured aggregate queue-limits of the classes, the system will drop packets on the inbound interface appearing as <input queue drops> and <ignores> in the interface statistic.

After the packet is processed by HQF it is placed in the final FIFO queue, the TX-Ring. The TX-Ring provides a final queue for the physical FIFO prior to transmit onto the wire.

When congestion occurs at the egress interface i.e. the TX-Ring is full, the CBWFQ queues begin to buffer traffic based on their per class queue-limit allocations. Ideally, these queues will reach their full capacity, then will begin to tail drop additional traffic in excess of the queue-limit configured. If however the router configuration for queue-limit exceeds the physical hold-queue limit, or the hold-queue exceeds the available buffering capabilities (private pools + public pools + additional I/O MEM RAM), traffic will be dropped prior to reaching the full buffer size configured. We will cover a few different ways to address the above situations, while detailing specific commands that require configuration.

Default Settings and Configuration Parameters

The current structure of buffering on the ISR platform is as follows:

1. Each physical interface is allocated 1000 packets of buffer (hold-queue)
2. Each newly instantiated class is allocated 64 packets of buffer (queue-limit)

3. The Tx-Ring is allocated **128/256** Packets of buffer (tx-ring-limit)

4. Upon consumption of physical buffers, interfaces request buffer allocation from the F/S buffer in I/O memory

The above values are configurable and viewable via the **show policy-map** interface**, show buffers, and show controllers** commands in Cisco IOS:

```
Router#show buffers

<snip>

Public particle pools:

F/S buffers, 1664 bytes (total 1536, permanent 1536):

     1024 in free list (256 min, 2048 max allowed)

     512 hits, 0 misses, 0 trims, 0 created

     0 failures (0 no memory)

     512 max cache size, 512 in cache

     0 hits in cache, 0 misses in cache

Normal buffers, 1676 bytes (total 3840, permanent 3840):

     3840 in free list (128 min, 4096 max allowed)

     0 hits, 0 misses, 0 trims, 0 created

     0 failures (0 no memory)

Private particle pools:

HQF buffers, 0 bytes (total 2000, permanent 2000):

     2000 in free list (500 min, 2000 max allowed)

     21 hits, 0 misses, 0 trims, 0 created

     0 failures (0 no memory)

IDS SM buffers, 240 bytes (total 128, permanent 128):

     0 in free list (0 min, 128 max allowed)

     128 hits, 0 fallbacks

     128 max cache size, 128 in cache

     0 hits in cache, 0 misses in cache

GigabitEthernet0/0 buffers, 1664 bytes (total 1024, permanent 1024):

     0 in free list (0 min, 1024 max allowed)

     1024 hits, 0 fallbacks

     1024 max cache size, 768 in cache

     333 hits in cache, 0 misses in cache

GigabitEthernet0/1 buffers, 1664 bytes (total 1024, permanent 1024):

     0 in free list (0 min, 1024 max allowed)

     1024 hits, 0 fallbacks
```

```
      1024 max cache size, 768 in cache

      343 hits in cache, 0 misses in cache<snip>
```

Keep in mind that the architecture of the memory handling has changed during IOS development. The latest IOS releases use the normal buffer as the backup pool for the private memory pools.

IOS Versions < 15.2T

The F/S buffers allocate a fast cache value (512) available directly, a permanent value (1536) which requires a fetch action, and a max value (512+1536=2048).

IOS Versions > 15.2T

The normal pool buffers allocate a fast cache value (256) available directly, a permanent value (3840) which requires a fetch action, and a max value (256+3840=4096).

Note: The "in cache" illustrates the available buffers, when this equals 0 the pool is empty.

The physical interface pool is broken down into rx-ring (256) + remainder (768) = max allowed (1024).

The tx-ring-limit is configurable in the interface configuration mode and can be verified by the <show controller command>:

```
Router(config)#int gi 0/0

Router(config-if)#tx

Router(config-if)#tx-ring-limit 200

Router(config-if)#

Router#sh controllers gi 0/0 | beg TX_COUNT

TX_COUNT: 0, TX_LIMITED: 1, TX_HEAD: 23, TX_TAIL: 23

TX_MIN_SIZE: 32, TX_MAX_SIZE: 512, TX_CURRENT_SIZE: 200

RX STARVES: 0

<snip>
```

This shows the current TX-Ring value on the interface after configuration modification. The default will be either 128/256.

```
Router#sh run po

Building configuration...

Current configuration : 198 bytes

!

policy-map buffering

 class cs4

  bandwidth percent 25

  queue-limit 128 packets

 class cs3
```

```
   bandwidth percent 10
 class class-default
   bandwidth percent 25
   queue-limit 256 packets
Router#sh policy-map int
 GigabitEthernet0/0
  Service-policy output: buffering

    Class-map: cs4 (match-all)
      0 packets, 0 bytes
      30 second offered rate 0 bps, drop rate 0 bps
      Match: ip dscp cs4 (32) af41 (34)
      Queueing
      queue limit 128 packets
      (queue depth/total drops/no-buffer drops) 0/0/0
      (pkts output/bytes output) 0/0
      bandwidth 25% (250000 kbps)


    Class-map: cs3 (match-all)
      0 packets, 0 bytes
      30 second offered rate 0 bps, drop rate 0 bps
      Match: ip dscp cs3 (24)
      Queueing
      queue limit 64 packets
      (queue depth/total drops/no-buffer drops) 0/0/0
      (pkts output/bytes output) 0/0
      bandwidth 10% (100000 kbps)

<snip>
```

Class cs4 has a modified queue-limit of 128 packets where class cs3 had a default of 64 packets.

Note: The aggregate of the queue-limits must be similar to the configured hold-queue limit (or the IOS default).

```
Router# sh int gi 0/0
GigabitEthernet0/0 is up, line protocol is up
  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d60 (bia 0026.cb27.6d60)
  Internet address is 10.70.1.1/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
```

```
    reliability 255/255, txload 1/255, rxload 1/255

  Encapsulation ARPA, loopback not set

  Keepalive set (10 sec)

  Full Duplex, 1Gbps, media type is RJ45

  output flow-control is XON, input flow-control is XON

  ARP type: ARPA, ARP Timeout 04:00:00

  Last input 00:00:07, output 00:00:00, output hang never

  Last clearing of "show interface" counters 5d04h

  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0

  Queueing strategy: Class-based queueing

  Output queue: 0/1000/0 (size/max total/drops)

<snip>
```

The show interface command illustrates the physical limit of 1000 packets of buffer as well as flow control negotiation. This is important to note as the previous ISR G1 (2800/3800) forced flow-control on, the current ISR G2 (2900/3900) negotiates flow control. Flow control will send pause frames to the source if oversubscription is occurring, and may assist in the buffering capabilities of the platform.

Configuring Buffer Parameters

When configuring the aggregate per class queue-limit below the hold-queue setting on the physical interface the ISR will show queue-depth, and total drops via the show policy-map interface command. In this example, the router is receiving 50Mbps of traffic spread across 3 DSCP values (EF, AF41, BE). The configurations shown are not best practice and should be used for demonstration purposes only. The policy and output is shown below:

```
Router#sh run policy-map

Building configuration...

Current configuration : 298 bytes

!

policy-map buffering

 class cs4

  bandwidth percent 25

  queue-limit 128 packets

 class cs3

  bandwidth percent 10

 class class-default

  bandwidth percent 25

  queue-limit 256 packets

policy-map buffering-parent
```

```
 class class-default
  shape average 9000000
  service-policy buffering
!
Router#sh policy-map int
 GigabitEthernet0/1
  Service-policy output: buffering-parent
    Class-map: class-default (match-any)
      2638885 packets, 3947650282 bytes
      30 second offered rate 54131000 bps, drop rate 45133000 bps
      Match: any
      Queueing
      queue limit 64 packets
      (queue depth/total drops/no-buffer drops) 386/2199701/0
      (pkts output/bytes output) 439197/656924874
      shape (average) cir 9000000, bc 36000, be 36000
      target shape rate 9000000
      Service-policy : buffering
        Class-map: cs4 (match-all)
          719512 packets, 1076389952 bytes
          30 second offered rate 14758000 bps, drop rate 10260000 bps
          Match: ip dscp cs4 (32) af41 (34)
          Queueing
          queue limit 128 packets
          (queue depth/total drops/no-buffer drops) 128/500019/0
          (pkts output/bytes output) 219493/328361528
          bandwidth 25% (2250 kbps)
```

The queue depth shown above correlates to the queue-limit, since the queue is being oversubscribed. Total drops demonstrates the tail drop on this specific queue.

In the following example, the aggregate queue-limit is configured above the hold-queue value, therefore the classes will not reach their full queue-limit capability.

```
Router#sh run policy-map
Building configuration...
Current configuration : 325 bytes
!
```

```
policy-map buffering
 class cs4
  bandwidth percent 25
  queue-limit 1000 packets
 class cs3
  bandwidth percent 10
  queue-limit 500 packets
 class class-default
  bandwidth percent 25
  queue-limit 256 packets
policy-map buffering-parent
 class class-default
  shape average 9000000
  service-policy buffering
```

The configuration above allows the per-class queue-limit to exceed the hold-queue (default 1000) on the interface. If this is the case, the classes are limited to the per interface hold-queue value, regardless of the per-class queue-limit configuration.

```
Router#sh policy-map int
 GigabitEthernet0/1
  Service-policy output: buffering-parent
    Class-map: class-default (match-any)
      531660 packets, 795340842 bytes
      30 second offered rate 53062000 bps, drop rate 43406000 bps
      Match: any
      Queueing
      queue limit 64 packets
      (queue depth/total drops/no-buffer drops) 1000/443256/0
      (pkts output/bytes output) 88404/132232420
      shape (average) cir 9000000, bc 36000, be 36000
      target shape rate 9000000
      Service-policy : buffering
        Class-map: cs4 (match-all)
          144961 packets, 216861656 bytes
          30 second offered rate 14472000 bps, drop rate 10185000 bps
          Match: ip dscp cs4 (32) af41 (34)
```

```
      Queueing
      queue limit 1000 packets
      (queue depth/total drops/no-buffer drops) 497/102468/0
      (pkts output/bytes output) 42493/63569528
      bandwidth 25% (2250 kbps)


    Class-map: cs3 (match-all)
      290041 packets, 433901336 bytes
      30 second offered rate 28951000 bps, drop rate 27139000 bps
      Match: ip dscp cs3 (24)
      Queueing
      queue limit 500 packets
      (queue depth/total drops/no-buffer drops) 500/273044/0
      (pkts output/bytes output) 16997/25427512
      bandwidth 10% (900 kbps)


    Class-map: class-default (match-any)
      96658 packets, 144577850 bytes
      30 second offered rate 9650000 bps, drop rate 6741000 bps
      Match: any
      Queueing
      queue limit 256 packets
      (queue depth/total drops/no-buffer drops) 0/67744/0
      (pkts output/bytes output) 28914/43235380
      bandwidth 25% (2250 kbps)
```

The cs4 class will never reach the 1000 packet buffer that is configured due to the limiting hold-queue default of 1000 packets. The parent shaper shows the queue-depth is at 1000 packets. The queue-limit value at the parent does not show the actual aggregate queue-limit and should be ignored.

Increasing the hold-queue resolves this issue.

```
Router# sh run int gi 0/1
Building configuration...
Current configuration : 155 bytes
!
interface GigabitEthernet0/1
 ip address 10.70.2.1 255.255.255.0
```

```
 load-interval 30
 duplex auto
 speed auto
 service-policy output buffering-parent
end
Router#config t
Enter configuration commands, one per line.  End with CNTL/Z.
Router(config)#int gi 0/1
Router(config-if)#hold-queue 2000 out
Router(config-if)#
Router#sh policy-map int
 GigabitEthernet0/1
  Service-policy output: buffering-parent
    Class-map: class-default (match-any)
      1268797 packets, 1898055018 bytes
      30 second offered rate 21984000 bps, drop rate 15025000 bps
      Match: any
      Queueing
      queue limit 64 packets
      (queue depth/total drops/no-buffer drops) 1584/1026584/0
      (pkts output/bytes output) 242212/362291821
      shape (average) cir 9000000, bc 36000, be 36000
      target shape rate 9000000
      Service-policy : buffering
        Class-map: cs4 (match-all)
          346767 packets, 518763432 bytes
          30 second offered rate 6128000 bps, drop rate 2278000 bps
          Match: ip dscp cs4 (32) af41 (34)
          Queueing
          queue limit 1000 packets
          (queue depth/total drops/no-buffer drops) 1000/232964/0
          (pkts output/bytes output) 113803/170249288
          bandwidth 25% (2250 kbps)


        Class-map: cs3 (match-all)
```

```
        688008 packets, 1029259968 bytes

        30 second offered rate 11265000 bps, drop rate 9736000 bps

        Match: ip dscp cs3 (24)

        Queueing

        queue limit 500 packets

        (queue depth/total drops/no-buffer drops) 500/642689/0

        (pkts output/bytes output) 45319/67797224

        bandwidth 10% (900 kbps)


      Class-map: class-default (match-any)

        234021 packets, 350030122 bytes

        30 second offered rate 4578000 bps, drop rate 946000 bps

        Match: any

        Queueing

        queue limit 256 packets

        (queue depth/total drops/no-buffer drops) 81/150931/0

        (pkts output/bytes output) 83090/124245309

        bandwidth 25% (2250 kbps)
```

From the example above, we see that the aggregate queue-limit is now less than the hold-queue limit. This allows each class queue to reach its per-class queue-limit configuration.

To demonstrate the limits of the platform memory, below, the max values via cli have been configured. In this case, the policy will run the router out of physical memory. This example is using fast switching buffers after the private interface pool is depleted. The subsequent example will demonstrate the use of normal buffers, using the same configuration and the same results.

```
Router#sh run po

Building configuration...

Current configuration : 330 bytes

!

policy-map buffering

 class cs4

  bandwidth percent 25

  queue-limit 32000 packets

 class cs3

  bandwidth percent 10

  queue-limit 32000 packets

 class class-default
```

```
    bandwidth percent 25
    queue-limit 32000 packets
policy-map buffering-parent
 class class-default
   shape average 9000000
   service-policy buffering
!
end
Router#sh run int gi 0/1
Building configuration...
Current configuration : 178 bytes
!
interface GigabitEthernet0/1
 ip address 10.70.2.1 255.255.255.0
 load-interval 30
 duplex auto
 speed auto
 service-policy output buffering-parent
 hold-queue 240000 out
end
```

The configuration again shows excess buffering, and a hold-queue size that will stretch the capacity of the platform.

```
Router#sh int gi 0/1
GigabitEthernet0/1 is up, line protocol is up
  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d61 (bia 0026.cb27.6d61)
  Internet address is 10.70.2.1/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
     reliability 255/255, txload 2/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1Gbps, media type is RJ45
  output flow-control is XON, input flow-control is XON
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:08, output 00:00:00, output hang never
```

```
Last clearing of "show interface" counters 00:18:25
Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 1400771
Queueing strategy: Class-based queueing
Output queue: 1725/240000/1400770 (size/max total/drops)
30 second input rate 0 bits/sec, 0 packets/sec
30 second output rate 8961000 bits/sec, 750 packets/sec
```

The egress port will begin to see output drops increase, due to the lack of physical buffer space available.

Public particle pools:

F/S buffers, 1664 bytes (total 1536, permanent 1536):

    1024 in free list (256 min, 2048 max allowed)

    512 hits, 0 misses, 0 trims, 0 created

    0 failures (0 no memory)

    512 max cache size, 0 in cache

    691929 hits in cache, 4363140 misses in cache

Normal buffers, 1676 bytes (total 3840, permanent 3840):

    3840 in free list (128 min, 4096 max allowed)

    0 hits, 0 misses, 0 trims, 0 created

    0 failures (0 no memory)

Private particle pools:

HQF buffers, 0 bytes (total 2000, permanent 2000):

    2000 in free list (500 min, 2000 max allowed)

    7977 hits, 0 misses, 0 trims, 0 created

    0 failures (0 no memory)

IDS SM buffers, 240 bytes (total 128, permanent 128):

    0 in free list (0 min, 128 max allowed)

    128 hits, 0 fallbacks

    128 max cache size, 128 in cache

    0 hits in cache, 0 misses in cache

GigabitEthernet0/0 buffers, 1664 bytes (total 1024, permanent 1024):

    0 in free list (0 min, 1024 max allowed)

1024 hits, 0 fallbacks

1024 max cache size, 0 in cache

7883664 hits in cache, 5055071 misses in cache

GigabitEthernet0/1 buffers, 1664 bytes (total 1024, permanent 1024):

0 in free list (0 min, 1024 max allowed)

1024 hits, 0 fallbacks

1024 max cache size, 540 in cache

38552 hits in cache, 0 misses in cache

From sh buffers we see the "in cache" equal to 0 on the ingress port, and the F/S buffers have also been consumed, therefore we will start to see ignores on ingress, as we reach for I/O memory buffers from the public pool.

At this point the egress interface has stopped dropping packets as "Total output drops", and the ingress interface is dropping packets as ignores, the buffers in the system are completely overrun, to the point that ingress traffic has no place to be buffered.

```
Router#sh int gi 0/0
GigabitEthernet0/0 is up, line protocol is up
  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d60 (bia 0026.cb27.6d60)
  Internet address is 10.70.1.1/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 2/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1Gbps, media type is RJ45
  output flow-control is XON, input flow-control is XON
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:03:40, output 00:00:01, output hang never
  Last clearing of "show interface" counters 00:05:40
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  30 second input rate 8949000 bits/sec, 748 packets/sec
  30 second output rate 0 bits/sec, 0 packets/sec
     254952 packets input, 381407120 bytes, 0 no buffer
     Received 1 broadcasts (0 IP multicasts)
```

```
        0 runts, 0 giants, 0 throttles

        0 input errors, 0 CRC, 0 frame, 0 overrun, 2822871 ignored

        0 watchdog, 1 multicast, 0 pause input

        41 packets output, 4343 bytes, 0 underruns

        0 output errors, 0 collisions, 0 interface resets

        0 unknown protocol drops

        0 babbles, 0 late collision, 0 deferred

        0 lost carrier, 0 no carrier, 0 pause output

        0 output buffer failures, 0 output buffers swapped out

Router#sh int gi 0/0

GigabitEthernet0/0 is up, line protocol is up

  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d60 (bia 0026.cb27.6d60)

  Internet address is 10.70.1.1/24

  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,

      reliability 255/255, txload 1/255, rxload 2/255

  Encapsulation ARPA, loopback not set

  Keepalive set (10 sec)

  Full Duplex, 1Gbps, media type is RJ45

  output flow-control is XON, input flow-control is XON

  ARP type: ARPA, ARP Timeout 04:00:00

  Last input 00:03:51, output 00:00:02, output hang never

  Last clearing of "show interface" counters 00:05:51

  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0

  Queueing strategy: fifo

  Output queue: 0/40 (size/max)

  30 second input rate 8970000 bits/sec, 751 packets/sec

  30 second output rate 0 bits/sec, 0 packets/sec

      263355 packets input, 393978008 bytes, 0 no buffer

      Received 1 broadcasts (0 IP multicasts)

      0 runts, 0 giants, 0 throttles

      0 input errors, 0 CRC, 0 frame, 0 overrun, 2915621 ignored

      0 watchdog, 1 multicast, 0 pause input

      42 packets output, 4403 bytes, 0 underruns

      0 output errors, 0 collisions, 0 interface resets
```

```
        0 unknown protocol drops

        0 babbles, 0 late collision, 0 deferred

        0 lost carrier, 0 no carrier, 0 pause output

        0 output buffer failures, 0 output buffers swapped out

Router#sh int gi 0/1

GigabitEthernet0/1 is up, line protocol is up

  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d61 (bia 0026.cb27.6d61)

  Internet address is 10.70.2.1/24

  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,

      reliability 255/255, txload 2/255, rxload 1/255

  Encapsulation ARPA, loopback not set

  Keepalive set (10 sec)

  Full Duplex, 1Gbps, media type is RJ45

  output flow-control is XON, input flow-control is XON

  ARP type: ARPA, ARP Timeout 04:00:00

  Last input 00:00:05, output 00:00:00, output hang never

  Last clearing of "show interface" counters 00:06:18

  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0

  Queueing strategy: Class-based queueing

  Output queue: 1784/240000/0 (size/max total/drops)

  30 second input rate 0 bits/sec, 0 packets/sec

  30 second output rate 8984000 bits/sec, 752 packets/sec

      19 packets input, 3233 bytes, 0 no buffer

      Received 19 broadcasts (0 IP multicasts)

      0 runts, 0 giants, 0 throttles

      0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

      0 watchdog, 19 multicast, 0 pause input

      283381 packets output, 423859898 bytes, 0 underruns

      0 output errors, 0 collisions, 0 interface resets

      13 unknown protocol drops

      0 babbles, 0 late collision, 0 deferred

      0 lost carrier, 0 no carrier, 0 pause output

      0 output buffer failures, 0 output buffers swapped out

Router#
```

After stopping the traffic, the public pool of memory shown via "show buffer" demonstrates free memory now available in the public pool.

Public particle pools:

F/S buffers, 1664 bytes (total 1536, permanent 1536):

    1024 in free list (256 min, 2048 max allowed)

    512 hits, 0 misses, 0 trims, 0 created

    0 failures (0 no memory)

    512 max cache size, 511 in cache

The Fast Switching (F/S) buffers are limited to the permanent value seen above, if one wanted to increase the value, it would allow the system to take more buffers out of the I/O memory, providing the ability to buffer a larger amount of traffic.

Increasing the fast switching buffers is not recommended, unless there is an explicit reason to do so. This configuration could run the entire platform out of I/O memory, causing the router to crash or worse. providing non-deterministic behavior.

```
Router(config)#buffers fastswitching permanent 4000
```

Public particle pools:

F/S buffers, 1664 bytes (total 4255, permanent **4000**):

    255 in free list (256 min, 2048 max allowed)

    3915 hits, 85 misses, 0 trims, 255 created

    0 failures (0 no memory)

    4000 max cache size, 3999 in cache

Using the same configuration as above with the latest image. the results are the same. Output drops are seen initially then when the public pool is depleted, ignores are seen on the ingress interface.

```
Router#sh int gi 0/1
GigabitEthernet0/1 is up, line protocol is up
  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d61 (bia 0026.cb27.6d61)
  Internet address is 10.70.2.1/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
     reliability 255/255, txload 2/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1Gbps, media type is RJ45
  output flow-control is XON, input flow-control is XON
  ARP type: ARPA, ARP Timeout 04:00:00
```

```
  Last input 00:00:18, output 00:00:00, output hang never

  Last clearing of "show interface" counters never

  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 40526

  Queueing strategy: Class-based queueing

  Output queue: 16457/240000/0 (size/max total/drops)

  30 second input rate 0 bits/sec, 0 packets/sec

  30 second output rate 11362000 bits/sec, 1423 packets/sec

    15 packets input, 2293 bytes, 0 no buffer

    Received 13 broadcasts (0 IP multicasts)

    0 runts, 0 giants, 0 throttles

    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

    0 watchdog, 13 multicast, 0 pause input

    1073555 packets output, 1069216441 bytes, 0 underruns

    0 output errors, 0 collisions, 0 interface resets

    9 unknown protocol drops

    0 babbles, 0 late collision, 0 deferred

    0 lost carrier, 0 no carrier, 0 pause output

    0 output buffer failures, 0 output buffers swapped out

Router#
```

The egress port will begin to see output drops increase due to the lack of physical buffer space available.

Also, a memory allocation failure begins to pop up on the screen:

```
*Apr  6 19:04:19.875: %SYS-2-MALLOCFAIL: Memory allocation of 65612 bytes failed
from 0x23C625FC, alignment 128

Pool: I/O  Free: 110880  Cause: Memory fragmentation

Alternate Pool: None  Free: 0  Cause: No Alternate pool

 -Process= "Pool Manager", ipl= 0, pid= 7

-Traceback= 23C4DB98z 23C5A024z 21502CE0z 23C5F5BCz 23C5F848z 23C3A8E8z 23C3A8CCz
```

Below via sh buffers, we see the "in cache" equal to 0 on the ingress port, and the normal buffers have also now been consumed, therefore we will start to see ignores on ingress as we reach for I/O memory buffers from the public pool.

Public particle pools:

F/S buffers, 1664 bytes (total 768, permanent 768):

    256 in free list (128 min, 1024 max allowed)

    512 hits, 0 misses, 0 trims, 0 created

0 failures (0 no memory)

512 max cache size, 512 in cache

0 hits in cache, 0 misses in cache

Normal buffers, 1676 bytes (total 15408, permanent 3840):

0 in free list (128 min, 4096 max allowed)

429286 hits, 8834250 misses, 18655 trims, 30223 created

8434667 failures (5977 no memory)

Private particle pools:

HQF Particle buffers, 0 bytes (total 2000, permanent 2000):

2000 in free list (500 min, 2000 max allowed)

72 hits, 0 misses, 0 trims, 0 created

0 failures (0 no memory)

IDS SM buffers, 240 bytes (total 128, permanent 128):

0 in free list (0 min, 128 max allowed)

128 hits, 0 fallbacks

128 max cache size, 128 in cache

0 hits in cache, 0 misses in cache

GigabitEthernet0/0 buffers, 1664 bytes (total 1024, permanent 1024):

76 in free list (0 min, 1024 max allowed)

1024 hits, 0 fallbacks

1024 max cache size, 0 in cache

858268 hits in cache, 8863962 misses in cache

GigabitEthernet0/1 buffers, 1664 bytes (total 1024, permanent 1024):

0 in free list (0 min, 1024 max allowed)

1024 hits, 0 fallbacks

1024 max cache size, 574 in cache

500 hits in cache, 0 misses in cache

At this point, the egress interface has stopped dropping packets as "Total output drops", and the ingress interface is dropping packets as ignores, the buffers in the system are completely overrun to the point that ingress traffic has no place to be buffered.

```
Router#sh int gi 0/0
```

```
GigabitEthernet0/0 is up, line protocol is up
  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d60 (bia 0026.cb27.6d60)
  Internet address is 10.70.1.1/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 2/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1Gbps, media type is RJ45
  output flow-control is XON, input flow-control is XON
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input never, output 00:00:02, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
  Queueing strategy: fifo
  Output queue: 0/40 (size/max)
  30 second input rate 8957000 bits/sec, 1125 packets/sec
  30 second output rate 0 bits/sec, 0 packets/sec
     1588785 packets input, 1582429860 bytes, 0 no buffer
     Received 0 broadcasts (0 IP multicasts)
     0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 0 overrun, 28735330 ignored
     0 watchdog, 0 multicast, 0 pause input
     88 packets output, 9937 bytes, 0 underruns
     0 output errors, 0 collisions, 0 interface resets
     0 unknown protocol drops
     0 babbles, 0 late collision, 0 deferred
     0 lost carrier, 0 no carrier, 187759 pause output
     0 output buffer failures, 0 output buffers swapped out

Router#sh int gi 0/1
GigabitEthernet0/1 is up, line protocol is up
  Hardware is CN Gigabit Ethernet, address is 0026.cb27.6d61 (bia 0026.cb27.6d61)
  Internet address is 10.70.2.1/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
     reliability 255/255, txload 2/255, rxload 1/255
```

```
     Encapsulation ARPA, loopback not set

     Keepalive set (10 sec)

     Full Duplex, 1Gbps, media type is RJ45

     output flow-control is XON, input flow-control is XON

     ARP type: ARPA, ARP Timeout 04:00:00

     Last input 00:00:03, output 00:00:00, output hang never

     Last clearing of "show interface" counters never

     Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 40526 (stays
the same)

     Queueing strategy: Class-based queueing

     Output queue: 16706/240000/0 (size/max total/drops)

     30 second input rate 0 bits/sec, 0 packets/sec

     30 second output rate 8953000 bits/sec, 1125 packets/sec

        39 packets input, 6533 bytes, 0 no buffer

        Received 37 broadcasts (0 IP multicasts)

        0 runts, 0 giants, 0 throttles

        0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

        0 watchdog, 37 multicast, 0 pause input

        1596486 packets output, 1589996211 bytes, 0 underruns

        0 output errors, 0 collisions, 0 interface resets

        25 unknown protocol drops

        0 babbles, 0 late collision, 0 deferred

        0 lost carrier, 0 no carrier, 0 pause output

        0 output buffer failures, 0 output buffers swapped out

Router#
```

After stopping the traffic, the public pool of memory shown via "show buffer" demonstrates free memory now available in the public pool. "In cache" value for normal buffers is back to 1024.

Public particle pools:

F/S buffers, 1664 bytes (total 768, permanent 768):

```
     256 in free list (128 min, 1024 max allowed)

     512 hits, 0 misses, 0 trims, 0 created

     0 failures (0 no memory)

     512 max cache size, 512 in cache

     0 hits in cache, 0 misses in cache
```

Normal buffers, 1676 bytes (total 4261, permanent 3840):

```
      4096 in free list (128 min, 4096 max allowed)
      858695 hits, 18959889 misses, 29802 trims, 30223 created    (all values stay
the same)
      18131040 failures (12851 no memory)
```

Private particle pools:

HQF Particle buffers, 0 bytes (total 2000, permanent 2000):

```
      2000 in free list (500 min, 2000 max allowed)
      147 hits, 0 misses, 0 trims, 0 created
      0 failures (0 no memory)
```

IDS SM buffers, 240 bytes (total 128, permanent 128):

```
      0 in free list (0 min, 128 max allowed)
      128 hits, 0 fallbacks
      128 max cache size, 128 in cache
      0 hits in cache, 0 misses in cache
```

GigabitEthernet0/0 buffers, 1664 bytes (total 1024, permanent 1024):

```
      334 in free list (0 min, 1024 max allowed)
      1024 hits, 0 fallbacks
      1024 max cache size, 1024 in cache
      894006 hits in cache, 18989735 misses in cache
```

GigabitEthernet0/1 buffers, 1664 bytes (total 1024, permanent 1024):

```
      0 in free list (0 min, 1024 max allowed)
      1024 hits, 0 fallbacks
      1024 max cache size, 343 in cache
      773 hits in cache, 0 misses in cache
```

Buffering Structure Summary

IOS Versions < 15.2T

The fast switching buffers are used to control the public pool of memory on the ISR G2. This memory is used only when the private interface buffer is completely used. The hold-queue determines the amount of egress buffer used by HQF at the physical interface, along with the per-class queue-limit commands. The aggregate per-class queue-limit should equate to the hold-queue, in order to allow buffering up to the per-class queue-limit.

It is not recommended, unless specifically required, to adjust the fast switching buffers. This memory is public to the platform, and is used not only to store packets, but for normal operation of the router. If the memory of the platform is exceeded, I/O memory exceed messages will be seen on the platform, which could lead to a router failure.

IOS Versions > 15.2T

Moving forward, the normal buffer pool, as opposed to the F/S pool, will be used to allocate buffer when physical buffer has been depleted. This is listed in the known issues section below as CSCtq27141.

Related Issues:

- CSCtj55834—2900 unable to use public buffers like the 2800
- CSCtq16226—Huge buffer values in 'buffers fastswitching permanent' causes crash
- CSCtw65356—Unaccounted packets with QOS/HQF configuration
- CSCtq27141—Buffer exhaustion occurs handling different in 2900 than 2800