Troubleshoot Catalyst 3850 Output Drops

Contents

Introduction

Prerequisites

Requirements

Components Used

Background Information

Buffer Allocation

Queue Structure

Tweak Buffer Allocation

Use Service Policy for Hardmax or Softmax Buffer Allocation

Example 1

Example 2

Example 3

Use Service Policy to Manually Change the Softmax Buffer Value

Example 1

Example 2

Case Study: Output Drops

Summary

Related Information

Introduction

This document describes the queue structure and buffers on the Catalyst 3650/3850 platform and provides examples on how output drops are mitigated.

Prerequisites

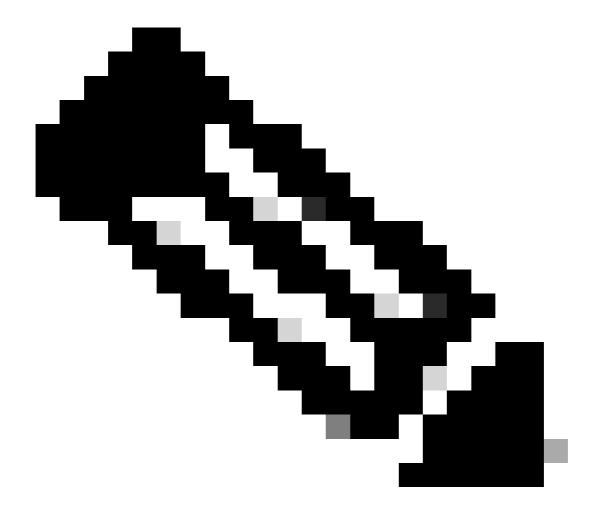
Requirements

Cisco recommends that you have basic knowledge of Quality of Service (QoS) on Catalyst platform.

Components Used

The information in this document is based on these software and hardware versions:

- Cisco Catalyst 3650/3850
- 03.07.04



Note: 16.x.x and later QoS CLI command changes are documented in this guide <u>Troubleshoot Output Drops on Catalyst 9000 Switches</u>. This document is Catalyst 9000 Series, but shares the same ASIC as the 3850. Use this guide for 3850 on 16.x.x or later Cisco IOS® XE versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Background Information

Output drops are generally a result of interface over-subscription caused by many to one or a 10gig to 1gig transfer. Interface buffers are a limited resource and can only absorb a burst up to a point after which packets can drop. You can get some cushion if you tune the buffers but it cannot guarantee a zero output drop scenario.

It is recommended to run 03.06 or 03.07's latest version to get appropriate buffer allocations due to some known bugs in older codes.

Buffer Allocation

Traditionally, buffers are statically allocated for each queue, and as you increase the number of queues the amount of reserved buffers decreases. This is inefficient and can deplete the number of buffers needed to handle frames for all queues. To get around that type of limitation, Catalyst 3650/3850 platform uses Hard buffers and Soft buffers.

- **Hard Buffers**: These are the minimum reserved buffers for specific queues. If a specific queue does not use the buffers, these are not available for other queues.
- **Soft Buffers**: These buffers are assigned to a queue but can be shared by other queues and interfaces if not used.

Default Buffer Allocation with No Service-policy Applied

The default buffer allocation for a 1GB port is 300 buffers and for a 10GB port, it is 1800 buffers (1 buffer = 256 bytes). The port can use up to 400% of the default allocated from common pool with default settings, which is 1200 buffers and 7200 buffers for 1 Gig interface and 10Gig interface respectively.

The default soft buffer limit is set to 400 (which is the max threshold). The threshold determines the maximum number of soft buffers that can be borrowed from the common pool.

Queue Structure

When no service-policy is applied, there are 2 default queues (queue 0 and queue 1). The queue-0 is used for control traffic (DSCP 32 or 48 or 56) and queue-1 is used for data traffic.

By default, queue 0 can be given 40% of the buffers that are available for the interface as its hard buffers. That is, 120 buffers are allocated for queue 0 in the context of 1G ports; 720 buffers in the context of 10G ports. The Softmax, the maximum soft buffers, for this queue is set to 480 (calculated as 400% of 120) for 1GB ports and 2880 for 10GB ports, where 400 is the default max threshold that is configured for any queue.

Queue 1 does not have any hard buffers allocated. The soft buffer value for queue-1 is calculated as 400% of the interface buffer that remains after it is allocated to queue-0. So, it is 400% of 180 for 1Gig interface and 400% of 1800 for a 10Gig interface.

The show command that can be used to see this allocation is show platform gos queue config <interface>.

For a 1Gig interface:

```
0
    1 5
120
 6
480
                   3 1440
 6
     320
    1 4
 1
0
                            3 1440
 7
                      180
     720
             480
 2
                    0 5
                                      0
                                          3
            0 5
                             0
                                0
                                            1440
 3
            0
               5
                    0 5
                             0
                                0
                                            1440
 4
    1
      4
            0 5
                    0 5
                            0
                               0
                                      0
                                         3
                                            1440
                    0 5
                           0 0
 5
    1 4
            0 5
                                      0
                                         3
                                            1440
    1
                     0 5
                             0 0
                                         3
            0 5
                                            1440
                                        3 1440
```

<<output omitted>>

For a 10Gig interface:

<#root>

3850#

show platform qos queue config tenGigabitEthernet 1/0/37

DATA Port:1 GPN:37 AFD:Disabled QoSMap:0 HW Queues: 8 - 15 DrainFast:Disabled PortSoftStart:2 - 6480

D ⁻	TS H	ard	max	Sof	tmax	Por	tSMin	G1b1	SMin	Port		
0	1	6	720	8	2880	7	1280	0	0	4	8640	
1	1	4	0	9	4320	8	1920	3	1080	4	8640	
2	1	4	0	5	0	5	0	0	0	4	8640	
3	1	4	0	5	0	5	0	0	0	4	8640	
4	1	4	0	5	0	5	0	0	0	4	8640	

<<output omitted>>

Hardmax or Hard Buffers is the amount of Buffer that is always reserved and available for this queue.

Softmax or Soft Buffers is the amount of buffer that can be borrowed from other queues or global pool. The total number of Softmax per 1Gig Interface is 1200 (400% of 300) and 7200 buffers if it is a 10Gig interface. When we apply a service-policy, there can be 1 extra queue created for "Class default" if not explicitly created. All the traffic that do not match under the previously defined classes fall under this queue. There cannot be any match statement under this queue.

Tweak Buffer Allocation

In order to tweak the buffers in 3650/3850 platform, attach a Service policy under the respective interface. You can tweak the Hardmax and Softmax buffer allocation with the service-policy.

Hard Buffer and Soft Buffer Calculations

This is how the system allocates Hardmax and Softmax for each queue:

```
Total Port buffer = 300 (1G) or 1800 (10G)
If there is a total of 5 queues (5 Classes), each queue gets 20% buffer by default.
```

Priority Queue

```
1Gig:
HardMax = Oper_Buff = 20% of 300 = 60.
qSoftMax = (Oper_Buff * Max_Threshold)/100=60*400/100=240

10Gig
HardMax = Oper_Buff = 20% of 1800 = 360
qsoftMax = (Oper_Buff * Max_Threshold)/100 = 360*400/100= 1440
```

Non-Priority Queue

```
1Gig:
HardMax = 0
qSoftMax = (Oper_Buffer*Max_Threshold)/100 = 300*20/100= 60. 400% of 60 = 240

10Gig:
HardMax = 0
qSoftMax = (Oper_Buffer*Max_Threshold)/100 = 1800*20/100= 360. 400% of 360 = 1440
```

Use Service Policy for Hardmax or Softmax Buffer Allocation

if a service-policy is applied, only the "Priority queue with level 1/2" gets the Hardmax. The next examples can help clarify the buffer allocation for specific service policy in 1Gig interface and 10Gig interface. With the default configuration that does not have any service policy applied, the queue-0 gets default Hardmax of 120 if the link is a 1Gig link and 720 buffers if the link is a 10Gig link.

```
show platform qos queue config tenGigabitEthernet 1/0/37
```

Example 1

When you apply a service-policy, if you do not configure a priority queue or if you do not set a priority queue level, there can be no Hardmax assigned to that queue.

For a 1Gig interface:

```
<#root>
policy-map MYPOL
class ONE
 priority percent 20
class TWO
 bandwidth percent 40
class THREE
 bandwidth percent 10
 class FOUR
 bandwidth percent 5
3850#
show run interface gig1/0/1
Current configuration: 67 bytes
interface GigabitEthernet1/0/1
service-policy output MYPOL
end
3800#
show platform qos queue config gigabitEthernet 1/0/1
DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175
 DrainFast:Disabled PortSoftStart:2 - 360
______
 DTS Hardmax Softmax PortSMin GlblSMin PortStEnd
```

```
___ _____
0
    1 4 0 8 240 7 160 3 60 4
                                         480
          0 8 240 7 160 3 60 4 480
1
   1 4
          0 8 240 7 160 3 60 4 480
2
           0 8 240 7
                         160 3 60 4 480
3
                         160 3
                 240 7
           0 8
                                  60 4
4
    1 4
                                          480
<<output omitted>>
!--- There are 5 classes present though you only created 4 classes.
!--- The 5th class is the default class.
!--- Each class represent a queue and the order in which it is shown is the order in which
!--- it is present in the running configuration when checking "show run | sec policy-map".
For a 10Gig interface:
<#root>
policy-map MYPOL
class ONE
 priority percent 20
class TWO
 bandwidth percent 40
class THREE
 bandwidth percent 10
class FOUR
 bandwidth percent 5
3850#
show run interface TenGigabitEthernet1/0/37
Current configuration: 67 bytes
interface TenGigabitEthernet1/0/37
service-policy output MYPOL
end
3850#
show platform qos queue config tenGigabitEthernet 1/0/40
DATA Port:2 GPN:40 AFD:Disabled QoSMap:1 HW Queues: 16 - 23
 DrainFast:Disabled PortSoftStart:4 - 2160
 DTS Hardmax Softmax PortSMin GlblSMin PortStEnd
             _____
         0 10 1440 9 640
    1 4
0
                                  360
                                      5 2880
1
  1 4
          0 10 1440 9 640 4 360 5 2880
          0 10 1440 9 640 4 360 5 2880
2
  1 4
3
          0 10 1440 9 640 4 360 5 2880
```

0 10 1440 9 640 4 360 5 2880

0 5 0 0 0 5 2880

Example 2

1 4

1 4 <<output omitted>>

0 5

```
When you apply priority level 1, the queue-0 gets 60 buffers as Hardmax.
For a 1Gig interface:
<#root>
 policy-map MYPOL
class ONE
priority level 1 percent 20
class TWO
 bandwidth percent 40
class THREE
 bandwidth percent 10
class FOUR
  bandwidth percent 5
3850#
show run interface gig1/0/1
Current configuration: 67 bytes
interface GigabitEthernet1/0/1
service-policy output MYPOL
BGL.L.13-3800-1#
show platform qos queue config gigabitEthernet 1/0/1
```

```
DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175 DrainFast:Disabled PortSoftStart:2 - 360
```

	DTS	 S H	 ardm	ax	Softmax		PortSMin	G1b1SI	 Min	PortStEnd	
0		1	6			-					
60											

8	2	40	7	1	60	0	0	4	480			
1	1	4		0	8	240	7	160	3	60	4	480
2	1	4		0	8	240	7	160	3	60	4	480
3	1	4		0	8	240	7	160	3	60	4	480
4	1	4		0	8	240	7	160	3	60	4	480

<<output omitted>>

For a 10Gig interface:

```
<#root>
```

policy-map MYPOL class ONE

```
priority level 1 percent 20
class TWO
 bandwidth percent 40
class THREE
 bandwidth percent 10
class FOUR
 bandwidth percent 5
3850#
show run interface Te1/0/37
Current configuration: 67 bytes
interface TenGigabitEthernet1/0/37
service-policy output MYPOL
3850_1#
show platform qos queue config tenGigabitEthernet 1/0/37
DATA Port:2 GPN:40 AFD:Disabled QoSMap:1 HW Queues: 16 - 23
 DrainFast:Disabled PortSoftStart:3 - 2160
 DTS Hardmax Softmax PortSMin GlblSMin PortStEnd
  ___ _____
   1 7
360
10
    1440 9
            640
                  0
                        0
                          5 2880
1
    1 4
            0 10 1440 9
                           640
                                    360
                                         5 2880
            0 10 1440 9
                                    360
 2
    1 4
                           640
                                         5 2880
 3
            0 10 1440 9
                           640 4
                                    360
                                         5 2880
            0 10 1440 9
                           640
                                         5 2880
    1 4
                                    360
            0 5
                    0 5
                           0
                                0
                                         5 2880
 5
    1 4
                                    0
<<output omitted>>
```

Example 3

For this example, one extra class is added. The total number of queues becomes 6. With 2 priority levels configured, each queue gets 51 buffers as Hardmax. The math is same as the previous example.

For 1Gig interface:

```
<#root>
policy-map MYPOL
class ONE

priority level 1 percent 20
class TWO
```

```
priority level 2 percent 10
```

class THREE
bandwidth percent 10
class FOUR
bandwidth percent 5
class FIVE
bandwidth percent 10

3850#

show run interface gigabitEthernet1/0/1

Current configuration : 67 bytes ! interface GigabitEthernet1/0/1 service-policy output MYPOL end

3850#

show platform qos queue config gigabitEthernet 1/0/1

```
DATA Port:16 GPN:10 AFD:Disabled QoSMap:1 HW Queues: 128 - 135 DrainFast:Disabled PortSoftStart:3 - 306
```

DTS Hardmax			Soft	Softmax		PortSMin		G1b1SMin		PortStEnd	
0	1	7									
51											
10 1	2 1	04 7	9	136	0	0	5	408			
51											
10	2	04	9	136	0	0	5	408			
2	1	4		0 10	204	9	136	4	51	5	408
3	1	4		0 10	204	9	136	4	51	5	408
4	1	4		0 11	192	10	128	5	48	5	408
5	1	4		0 11	192	10	128	5	48	5	408
6	1	4		0 5	0	5	0	0	0	5	408

For a 10Gig interface:

<<output omitted>>

<#root>

policy-map MYPOL
 class ONE

priority level 1 percent 20

class TWO

```
priority level 2 percent 10
class THREE
 bandwidth percent 10
class FOUR
 bandwidth percent 5
class FIVE
 bandwidth percent 10
3850#
show run interface Te1/0/37
Current configuration: 67 bytes
interface TenGigabitEthernet1/0/37
service-policy output MYPOL
end
3850_2#
show platform qos queue config tenGigabitEthernet 1/0/37
DATA Port:2 GPN:40 AFD:Disabled QoSMap:1 HW Queues: 16 - 23
 DrainFast:Disabled PortSoftStart:4 - 1836
_____
 DTS Hardmax Softmax PortSMin GlblSMin PortStEnd
   1 8
306
12 1224 11 544 0 0 6 2448
1
    1 8
306
12
    1224 11 544
                0
                      0
                         6 2448
            0 12 1224 11
                          544 6
                                  306 6 2448
2
3
            0 12 1224 11 544
                               6
                                  306 6 2448
    1 4
            0 13 1152 12
                          512 7
    1 4
                                  288
                                      6 2448
4
                         512 7
            0 13 1152 12
    1 4
                                  288
                                       6
                                          2448
    1 4
            0 5
                 0 5 0 0
                                 0
                                       6 2448
<<output omitted>>
```



Note: There can be less buffers allocated to few queues. This is expected as the values that cannot fit into Softmax calculation for priority queue and non-priority queue while certain configurations are combined.

In summary, the more queues you create, the less buffers each queue gets in terms of Hardmax and Softmax (as Hardmax is also dependant on Softmax value).

From 3.6.3 or 3.7.2, the maximum value for Softmax can be modified with the CLI command qos queuesoftmax-multiplier 1200 with 100 as the default value. If configured as 1200, the Softmax for non-priority queues and non-primary priority queue (!=level 1) are multiplied by 12 from their default values. This command would take effect only on the ports where a policy-map is attached. It is also not applicable for priority

queue level 1.

Use Service Policy to Manually Change the Softmax Buffer Value

Example 1

This is the service policy configuration and the correspondent buffer allocation:

```
<#root>
policy-map TEST_POLICY
 class ONE
  priority level 1 percent 40
 class TWO
  bandwidth percent 40
 class THREE
  bandwidth percent 10
3850#
show run interface gigabitEthernet1/0/1
Current configuration : 67 bytes
interface GigabitEthernet1/0/1
 service-policy output TEST_POLICY
3850#
show platform qos queue config gigabit\texttt{E}thernet 1/0/1
DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175
  DrainFast:Disabled PortSoftStart:2 - 450
  DTS Hardmax Softmax PortSMin GlblSMin PortStEnd
 0 1 6 75 8 300 7 200 0 0 4 600

    1
    4
    0
    8
    300
    7
    200
    3
    75
    4
    600

    1
    4
    0
    8
    300
    7
    200
    3
    75
    4
    600

    1
    4
    0
    8
    300
    7
    200
    3
    75
    4
    600

    1
    4
    0
    8
    300
    7
    200
    3
    75
    4
    600

 1
<<output omitted>>
```

The buffers are equally split across the queues. If you use the bandwidth command only the weight is changed for every queue and how the scheduler acts on it.

To tweak the Softmax value, you have to use the queue-buffer ratio command under the respective class.

```
<#root>
policy-map TEST_POLICY
class ONE
```

```
priority level 1 percent 40
class TWO
bandwidth percent 40
```

queue-buffers ratio 50

<----class THREE
bandwidth percent 10
class FOUR
bandwidth percent 5</pre>

These are the new buffer allocations.

For 1Gig interface:

<#root>

3850#

show platform qos queue config gigabit
Ethernet 1/0/1

DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175 DrainFast:Disabled PortSoftStart:0 - 900

DTS Hardmax				Softmax		PortSMin		G1b1SMin		Port	StEnd	
0	1	6	39	8	156	7	104	0	0	0	1200	
1	1	4	0	9								
600												
8	4	00	3	150	0	1200						
2	1	4	0	8	156	7	104	4	39	0	1200	
3	1	4	0	10	144	9	96	5	36	0	1200	
4	1	4	0	10	144	9	96	5	36	0	1200	

The queue-1 gets 50% of the soft buffer, that is, 600 buffers. The rest of the buffers are allocated to the other queues as per the algorithm.

Similar output for a 10Gig interface is:

<#root>

3850#

show platform qos queue config tenGigabitEthernet 1/0/37

DATA Port:2 GPN:40 AFD:Disabled QoSMap:1 HW Queues: 16 - 23

DrainFast:Disabled PortSoftStart:4 - 1836

DTS Hardmax Softmax PortSMin GlblSMin PortStEnd

0 1 7 234 10 936 9 416 0 0 5 7200

1 1 4 0 11

10	16	00	4	900	5	720	0				
2	1	4	0	10	936	9	416	5	234	5	7200
3	1	4	0	4	864	11	384	1	216	5	7200
4	1	4	0	4	864	11	384	1	216	5	7200
5	1	4	0	5	0	5	0	0	0	5	7200
< <output omitted="">></output>											



Note: There can be fewer buffers allocated to a few queues. This is expected as the values cannot fit into Softmax calculation for priority queue and non-priority queue when certain configurations are combined. There is an internal algorithm which takes care of it.

Example 2

Allocate all of the Softmax buffer to the single default queue.

You can see less buffers allocated to a few queues. This is expected as the values cannot fit into the Softmax calculation for priority queue and non-priority queue when certain combinations are configured. There is an internal algorithm which takes care of it. There can be less buffers allocated to few queues. This is expected as the values cannot fit into Softmax calculation for priority queue and non-priority queue when certain combinations are configured. There is an internal algorithm which take care of it.

```
policy-map NODROP
 class class-default
 bandwidth percent 100
 queue-buffers ratio 100
```

The QoS configuration results are:

```
<#root>
```

3850#

show platform qos queue config GigabitEthernet 1/1/1

DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175 DrainFast:Disabled PortSoftStart:0 - 900

DT	 S H	ardr	nax	Sof	tmax	Por	tSMin	G1b1:	SMin	Port	StEnd
0	 1	4	 0	8	1200	7	800	3	300	2	2400
1	1	4	0	5	0	5	0	0	0	2	2400

There is no Hardmax buffer since the policy is applied to an interface and it does not have any priority queue with "level" set. As soon as you apply the policy-map, the second queue is disabled and that leaves only 1 queue in the system.

The caveat here is that all packets use this single queue (that includes the control packets like

OSPF/EIGRP/STP). When there is congestion (broadcast storm), this can easily cause network disruption. This also occurs if you have other classes defined the control packets that match.

Case Study: Output Drops

For this test, IXIA traffic generator is connected to 1Gig interface and the egress port is 100Mbps interface. This is a 1Gbps to 100Mbps connection and a burst of 1 Gig of packets are sent for 1 second. This can cause output drop on the egress 100mbps interface. With the default configuration (no service-policy applied), this is the number of output drops after 1 is sent:

```
<#root>
3850#
show interfaces gig1/0/1 | in output drop
  Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops:
497000
```

These drops are seen in Th2, which is the default threshold. By the default, the system can use max threshold as drop threshold which is Drop-Th2.

After that, configure this service-policy to tweak the buffer:

```
<#root>

policy-map TEST_POLICY
  class class-default
  bandwidth percent 100
  queue-buffers ratio 100

3850#
```

show run interface gigabitEthernet1/0/1

```
Current configuration : 67 bytes ! interface GigabitEthernet1/0/1 service-policy output TEST_POLICY end
```

3850#

show platform qos queue config gigabitEthernet 2/0/1

DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175 DrainFast:Disabled PortSoftStart:0 - 900

DTS Hardmax Softmax PortSMin GlblSMin PortStEnd
--- -----0 1 4 0 8

7 800 3 300 2 2400 <-- queue 0 gets all the buffer.

3850#

show interfaces gigabitEthernet1/0/1 | include output drop

Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops:

385064

The drops reduced from 497000 to 385064 for the same traffic burst. Yet, there are still drops. After that, configure qos queue-softmax-multiplier 1200 global config command.

<#root>

3850#

show platform qos queue config gigabitEthernet 1/0/1

DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175 DrainFast:Disabled PortSoftStart:0 - 900

3850#

show interfaces gigabitEthernet1/0/1 | in output drop

Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: 0

The Softmax for queue-0 can go up to 10,000 buffers and as a result, the drops are Zero.



Note: This kind of scenario is not possible as other interfaces can also use the buffer, but, this can definitely help to reduce the packet drops to a certain level.

The maximum soft buffer available for an interface can be increased with this command however, you must also keep in mind that this is available only if no other interface uses these buffers.

Summary

- 1. When you create more queues, you get less buffer for each queue.
- 2. The total number of buffers available can be increased with qos queue-softmax-multiplier <value>command.
- 3. If you define only 1 class-default, in order to tweak the buffer, all the traffic falls under the single queue (that includes control packets). Be advised that when all traffic is put in one queue, there is no classification between control and data traffic and during time of congestion, control traffic could get dropped. So, it is recommended to create at least 1 other class for control traffic. CPU generated control-packets always go to the first priority queue even if not matched in the class-map. If there is no priority queue configured, it would go to the first queue of the interface, which is queue-0.
- 4. Prior to Cisco bug ID CSCuu14019, interfaces wont display "output drop" counters, you have to execute show platform gos queue stats output to check for drops.
- 5. An enhancement request, Cisco bug ID CSCuz86625, was submitted to let us configure soft-max multiplier without the use of any service-policy.(Resolved in 3.6.6 and above)

Related Information

• Cisco Technical Support & Downloads