

Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Components Used](#)

[Buffer Allocation](#)

[Queue Structure](#)

[Tweaking Buffer Allocation](#)

[Using Service policy for Hardmax or Softmax buffer allocation](#)

[Example 1](#)

[Example 2](#)

[Example 3](#)

[Using Service policy for manually changing the Softmax buffer value:](#)

[Example 1](#)

[Example 2](#)

[Case Study: Output Drops](#)

[Summary](#)

Introduction

The document will help to provide an understanding of the queue structure and buffers on the Catalyst 3650/3850 platform. It also provides example on how output drops can be mitigated to a certain extent.

Output drops are generally a result of interface oversubscription 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 will drop. Tuning the buffers can give you some cushion 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.

Prerequisites

Requirements

Cisco recommends that you have basic knowledge of QoS on Catalyst platform.

Components Used

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

- Cisco Catalyst 3850
- 03.07.04

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 was inefficient and could lead to not having enough buffers 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: This is the minimum reserved buffer for specific queues. If a specific queue does not use the buffers, it is not available for other queues.

Soft Buffers: These buffers are assigned to a queue but can be shared by other queues and interfaces if it's not being used.

Default Buffer Allocation with no service-policy is 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 upto 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 would determine 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 will be given 40% of the buffers that are available for the interface as its hard buffers. i.e. 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's remaining buffer after being 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 qos queue config <interface>'.

For a 1Gig interface,

For a 10Gig interface,

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 will be 1 extra queue created for "Class default" if not explicitly created. All the traffic that are not matching under the previously defined classes fall under this queue. There can not be any match statement under this queue.

Tweaking Buffer Allocation

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

Hard buffer and Soft buffer calculations:

This is how the system allocates softmax and hardmax 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

Using 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. Below examples will help clarify the buffer allocation for specific service policy in 1Gig interface and 10Gig interface.

As we know, With default configuration where you have not applied any service policy, 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.

Example 1

While applying a service-policy, if you don't configure a priority queue or if you don't set a priority

queue level, there will be no hardmax assigned to that queue

For a 1Gig interface:

For a 10Gig interface:

Example 2

When you apply "priority level 1", the queue-0 gets 60 buffers as Hardmax. There is a little math behind this and it was explained in the SoftMax and HardMax calculation section earlier.

For a 1Gig interface:

For a 10Gig interface:

Example 3

For this 3rd Example, i am going to add one extra class. now 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:

For a 10Gig interface:

Note: Sometimes you may see less buffers allocated to few queues. This is expected as the values cant fit into Softmax calculation for priority queue and non-priority queue during certain combination of configurations.

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).

Note: Starting from 3.6.3 or 3.7.2, the maximum value for softmax can be modified using a CLI command: "qos queue-softmax-multiplier 1200", with 100 being 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.

Using Service policy for manually changing the Softmax buffer value:

Example 1

The service policy configuration and the corresponding buffer allocation is shown below

The buffers are equally split across the queues. Using bandwidth command will only change the weight for every queue and correspondingly how the scheduler will act on it.

To tweak the softmax value, you have to use "Queue-buffer ratio" command under the respective class.

The new buffer allocations are:

For 1gig interface:

Now, queue-1 gets 50% of the soft buffer, i.e: 600 buffers. the remaining buffers are allocated to the other queues as per the algorithm.

Similar output for a 10-gig interface is:

Note: sometimes you may see less buffers allocated to few queues. This is expected as the values cant fit into Softmax calculation for priority queue and non-priority queue during certain combination of configurations. There is an internal algorithm which takes care of it.

Example 2

Allocating all the softmax buffer to the single default queue

The QOS config results are as follows:

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 2nd queue gets disabled leaving only 1 queue in the system.

The caveat here is that all packets are going to use this single queue (including the control packets like OSPF/EIGRP/STP).

During the time of congestion (broadcast storm etc), this may easily cause network disruption.

This is true if you have other classes defined but not matching the control packets.

Case Study: Output Drops

For this test, IXIA traffic genertor 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 will cause output drop on the egress 100mbps interface.

With the default config (no service-policy applied), the number of output drops after sending 1 stream is shown below

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

After configuring the following service-policy to tweak the buffer,

The drops reduced from 497000 to 385064 for a same traffic burst. yet, there are still drops.

After configuring "qos queue-softmax-multiplier 1200" global config command.

Now, the softmax for queue-0 can go upto 10,000 buffers and as a result, the drops are Zero.

Note: In real life, this kind of scenario may not be possible as other interfaces may also use the buffer, but, this can definately help in reducing the packet drops to a certain level.

The maximum soft buffer available for an interface can be increased using this command however, you should also keep in mind that this is available only if no other interface is using 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 using "qos queue-softmax-multiplier X" command.
3. If you define only 1 class-default, in order to tweak the buffer, all the traffic falls under the single queue (including 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 will 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 [CSCuu14019](#) , interfaces wont display "output drop" counters. you have to execute "show platform qos queue stats" output to check for drops.
5. An enhancement request, [CSCuz86625](#) , has been submitted to let us configure soft-max multiplier without using any service-policy.(Resolved in 3.6.6 and above)