Catalyst 3850輸出捨棄疑難排解

目錄

簡介 必要條件 需求 採用元件 背景資訊 緩衝區分配 隊列結構 微調緩衝區分配 將服務策略用於Hardmax或Softmax緩衝區分配 範例 1 範例 2 範例 3 使用服務策略手動更改Softmax緩衝區值 範例 1 範例 2 <u>案例研究:輸出下降</u> 摘要 相關資訊

簡介

本檔案將說明Catalyst 3650/3850平台上的佇列結構和緩衝區,並提供有關如何減輕輸出捨棄的範例。

必要條件

需求

思科建議您瞭解Catalyst平台上的服務品質(QoS)的基本知識。

採用元件

本文中的資訊係根據以下軟體和硬體版本:

- Cisco Catalyst 3650/3850
- 03.07.04

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除(預設))的組態來啟動。如果您的網路運作中,請確保您瞭解任何指令可能造成的影響。

背景資訊

輸出丟棄通常是由於多對一或10gig到1gig傳輸引起的介面超訂用所致。介面緩衝區是有限的資源 ,只能吸收最多到資料包可以丟棄的點的突發。如果調整緩衝區,您可以獲得一些緩衝,但它不能 保證出現零輸出丟棄情況。

建議運行03.06或03.07的最新版本,以獲取合適的緩衝區分配,因為舊代碼中存在一些已知的錯誤 。

緩衝區分配

傳統上,會為每個隊列靜態分配緩衝區,並且隨著隊列數量的增加,預留緩衝區的數量也會減少。 這樣效率低下,並且會耗盡所有隊列處理幀所需的緩衝區數量。為繞過此類限制,Catalyst 3650/3850平台使用硬緩衝區和軟緩衝區。

- •硬緩衝區:這些是特定隊列的最小保留緩衝區。如果特定隊列不使用緩衝區,則這些緩衝區對 其他隊列不可用。
- 軟緩衝:這些緩衝區已分配給隊列,但是如果不使用,則可以由其他隊列和介面共用。

未應用服務策略的預設緩衝區分配

1GB埠的預設緩衝區分配是300個緩衝區,而10GB埠的預設緩衝區分配是1800個緩衝區(1個緩衝 區= 256位元組)。該埠最多可以使用從具有預設設定的公共池分配的400%的預設值,即分別為1 Gig介面和10 Gig介面分配1200個緩衝區和7200個緩衝區。

預設軟緩衝區限制設定為400(即最大閾值)。閾值確定可以從公用池借用的軟緩衝區的最大數量 。

隊列結構

不應用服務策略時,有2個預設隊列(隊列0和隊列1)。queue-0用於控制流量(DSCP 32或48或 56),queue-1用於資料流量。

預設情況下,可為隊列0提供40%的介面可用緩衝區作為其硬緩衝區。也就是說,在1G埠的情況下 為隊列0分配了120個緩衝區;在10G埠的情況下分配了720個緩衝區。1GB埠的Softmax(最大軟緩 衝區)設定為480(計算為120的400%),10GB埠的Softmax(最大軟緩衝區)設定為2880,其中 400是為任何隊列配置的預設最大閾值。

隊列1未分配任何硬緩衝區。queue-1的軟緩衝區值計算為分配給queue-0後剩餘介面緩衝區的 400%。因此,1Gig介面為180%的400%,而10Gig介面為1800的400%。

其 show可用於檢視此分配的命令是 show platform qos queue config.

對於1Gig介面,

3850#show platform gos queue config gigabitEthernet 1/0/1
DATA Port:20 GPN:66 AFD:Disabled QoSMap:0 HW Queues: 160 - 167
DrainFast:Disabled PortSoftStart:1 - 1080

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

<<output omitted>>

對於10Gig介面,

3850#show platform gos queue config tenGigabitEthernet 1/0/37	
DATA Port:1 GPN:37 AFD:Disabled QoSMap:0 HW Queues: 8 - 15	
DrainFast:Disabled PortSoftStart:2 - 6480	

DIS Harullax			llldX	SOL	LIIIAX	POLCSMIN		GIDI	SMITH	POLISIENA		
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是始終為此隊列保留和可用的Buffer量。

Softmax or Soft Buffers是可從其他隊列或全域性池借用的緩衝區量。每個1Gig介面的Softmax總數 為1200(300的400%),如果是10Gig介面,則有7200個緩衝區。 應用服務策略時,如果沒有顯 式建立,則可以為「Class default」建立1個額外隊列。所有在前面定義的類下不匹配的流量都屬於 此隊列。此隊列下不能有任何match語句。

微調緩衝區分配

為了調整3650/3850平台中的緩衝區,請在各自的介面下附加服務原則。您可以使用servicepolicy調整Hardmax和Softmax緩衝區分配。

硬緩衝區和軟緩衝區的計算

這是系統為每個隊列分配Hardmax和Softmax的方式:

總埠緩衝區= 300(1G)或1800(10G) 如果共有5個隊列(5類),則預設情況下每個隊列將獲得20%的緩衝區。

優先順序隊列

1Gig: HardMax = Oper_Buff = 20% of 300 = 60。 gSoftMax =(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

非優先順序佇列

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

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

將服務策略用於Hardmax或Softmax緩衝區分配

如果應用服務策略,則只有「級別為1/2的優先順序隊列」才能獲得Hardmax。下面的示例可幫助澄 清1Gig介面和10Gig介面中特定服務策略的緩衝區分配。在預設配置未應用任何服務策略的情況下 ,如果鏈路為1Gig鏈路,則queue-0的預設硬體最大值為120;如果鏈路為10Gig鏈路,則預設硬體 最大值為720。

3850 DAT# Di	3850#show platform qos queue config gigabitEthernet 1/0/1 DATA Port:0 GPN:119 AFD:Disabled QoSMap:0 HW Queues: 0 - 7 DrainFast:Disabled PortSoftStart:1 - 1080													
D	DTS Hardmax Softmax PortSMin GlblSMin PortStEnd													
0	1	5	120	6	480	6	320	0	0	3	1440			
1	1	4	0	7	720	3	480	2	180	3	1440			
2	1	4	0	5	0	5	0	0	0	3	1440			
				-										

<<output omitted>>

3850#show platform gos queue config tenGigabitEthernet 1/0/37

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

	DTS Hardmax				Soft	max	Port	SMin	Glbl	SMin :	Ports	StEnd
	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
<	<out< td=""><td>put</td><td>om:</td><td>itte</td><td>d>></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></out<>	put	om:	itte	d>>							

範例 1

應用服務策略時,如果未配置優先順序隊列或未設定優先順序隊列級別,則不能為該隊列分配硬最 大值。

對於1Gig介面:

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	Ha	rdmax	S	oftr	nax P	SMin Gl	blsM	lin Po	rtSt	End	
			-								
0	1	4	0	8	240	7	160	3	60	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>>

!--- 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".

_ _

對於10Gig介面:

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	Har	rdmax	ç	Softn	nax Po	ortSN	4in Glk	olSM:	in Poi	rtSt	End
			-								
0	1	4	0	10	1440	9	640	4	360	5	2880
1	1	4	0	10	1440	9	640	4	360	5	2880
2	1	4	0	10	1440	9	640	4	360	5	2880
3	1	4	0	10	1440	9	640	4	360	5	2880
4	1	4	0	10	1440	9	640	4	360	5	2880

5	1	4	0	5	0	5	0	0	0	5	2880	< <output< th=""><th>omitted>></th></output<>	omitted>>
---	---	---	---	---	---	---	---	---	---	---	------	---	-----------

範例 2

應用優先順序級別1時, queue-0獲得60個緩衝區作為Hardmax。

對於1Gig介面:

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 end

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	Ha	rdmax		Softr	nax	Port	tSMin	Glbl	SMin 1	Ports	StEnd
			_								
0	1	6	60	8	240	7	160	0 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>>

對於10Gig介面:

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
end
```

DATA Port:2 GPN:40 AFD:Disabled QoSMap:1 HW Queues: 16 - 23 DrainFast:Disabled PortSoftStart:3 - 2160

DI	rs h	ard	max	Sof	tmax	PortSMin		GlblSMin		PortStEnd	
0	1	7	360	10	1440	9	640	0	0	5	2880
1	1	4	0	10	1440	9	640	4	360	5	2880
2	1	4	0	10	1440	9	640	4	360	5	2880
3	1	4	0	10	1440	9	640	4	360	5	2880
4	1	4	0	10	1440	9	640	4	360	5	2880
5 1 4 0			5	0	5	0	0	0	5	2880	
<<01	ıtpu	t o	mitted	1>>							

範例 3

在此示例中,新增了一個額外的類。隊列總數變為6。配置了2個優先順序後,每個隊列將獲得51個 緩衝區作為Hardmax。此數學運算與上一個範例相同。

對於1Gig介面:

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	5 Ha	rdmax	5	Soft	max 1	Port	SMin	GlblSMin PortStEnd			
0	1	7	51	10	204	9	136	0	0	5	408
1	1	7	51	10	204	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
<<01	output omitted>>										

對於10Gig介面:

```
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 gos queue config tenGigabitEthernet 1/0/37

DATA Port:2 GPN:40 AFD:Disabled QoSMap:1 HW Queues: 16 - 23 DrainFast:Disabled PortSoftStart:4 - 1836

D	DTS Hardmax		lmax	Softmax		PortSMin		n GlblSMin		PortStEnd	
-											
0	1	8	306	12	1224	11	544	0	0	6	2448
1	1	8	306	12	1224	11	544	0	0	6	2448
2	1	4	0	12	1224	11	544	6	306	6	2448
3	1	4	0	12	1224	11	544	6	306	6	2448
4	1	4	0	13	1152	12	512	7	288	6	2448
5	1	4	0	13	1152	12	512	7	288	6	2448
6	1	4	0	5	0	5	0	0	0	6	2448
				-							

<<output omitted>>

注意:分配給少數隊列的緩衝區可能更少。當組合某些配置時,這是優先順序隊列和非優先順 序隊列的Softmax計算無法滿足的值。

總之,您建立的隊列越多,每個隊列在Hardmax和Softmax方面獲得的緩衝區就越少(因為 Hardmax也取決於Softmax值)。

從3.6.3或3.7.2開始,可以使用CLI命令修改Softmax的最大值 qos queue-softmax-multiplier 1200 預設值為 100。如果配置為1200,則非優先順序隊列和非主優先順序隊列(!=1級)的Softmax將從其預設值乘 以12。此命令僅在連線了策略對映的埠上生效。它也不適用於優先順序隊列級別1。

使用服務策略手動更改Softmax緩衝區值

範例 1

這是服務策略配置和往來行緩衝區分配:

```
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

end
```

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:2 - 450

DTS Hardmax			ax S	Softmax			PortSMin GlblSMin				PortStEnd		
0	1	6	75	8	300	7	200	0	0	4	600		
1	1	4	0	8	300	7	200	3	75	4	600		
2	1	4	0	8	300	7	200	3	75	4	600		
3	1	4	0	8	300	7	200	3	75	4	600		
< <output omitted="">></output>													

緩衝區在隊列中平均分配。如果使用bandwidth命令,則僅更改每個隊列的權重以及排程程式對其執 行的方式。

要調整Softmax值,必須使用 queue-buffer ratio 命令。

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 這些是新的緩衝區分配。

對於1Gig介面:

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 _____ DTS Hardmax Softmax PortSMin GlblSMin PortStEnd --- ----- ------ ------ ------0 16 39 8 156 7 104 0 0 0 1200 1 4 0 9 **600** 8 400 3 150 0 1200 1 0 8 156 7 104 4 39 0 1200 1 4 2 0 10 144 9 96 5 36 0 1200 1 4 3 96 144 9 5 4 1 4 0 10 36 0 1200

queue-1獲得50%的軟緩衝區,即600個緩衝區。根據演算法,其餘的緩衝區會分配到其他隊列。

10Gig介面的類似輸出為:

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	3600	10	1600	4	900	5	7200	
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>>

注意:分配給少數隊列的緩衝區可以更少。這是預期值,因為當組合某些配置時,值無法適合 優先順序隊列和非優先順序隊列的Softmax計算。有一個內部演算法可以處理它。

範例 2

將所有Softmax緩衝區分配給單個預設隊列。

您可以看到分配給幾個隊列的緩衝區更少。這是預期值,因為當配置了某些組合時,這些值無法適 合優先順序隊列和非優先順序隊列的Softmax計算。有一個內部演算法可以處理它。分配給少數隊 列的緩衝區可以更少。這是預期值,因為當配置了某些組合時,這些值無法適合優先順序隊列和非 優先順序隊列的Softmax計算。有一個內部演算法可以處理它。

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

QoS配置結果如下:

3850#show platform gos queue config GigabitEthernet 1/1/1
DATA Port:21 GPN:65 AFD:Disabled QoSMap:1 HW Queues: 168 - 175
DrainFast:Disabled PortSoftStart:0 - 900

DTS Hardmax		Softmax		PortSMin		GlblSMin		PortStEnd					
				-									
() 1	L	4	0	8	1200	7	800	3	300	2	2400	
1	1 1	L	4	0	5	0	5	0	0	0	2	2400	

沒有Hardmax緩衝區,因為該策略應用於介面,並且它沒有任何設定了「level」的優先順序隊列。 應用策略對映後,第二個隊列即被禁用,這僅會在系統中保留1個隊列。

這裡需要注意的是,所有資料包都使用此單一隊列(包括控制資料包,如OSPF/EIGRP/STP)。 發生擁塞(廣播風暴)時,很容易造成網路中斷。 如果其它類定義了匹配的控制資料包,也會發生 這種情況。

案例研究:輸出下降

在本測試中,IXIA流量發生器連線到1Gig介面,輸出埠為100Mbps介面。這是一個1Gbps到

100Mbps的連線,並且一組1 Gig的資料包會被傳送1秒。這可能會導致輸出100mbps介面的輸出下 降。使用預設配置(未應用服務策略)時,這是傳送1後的輸出丟棄數:

3850#show interfaces gig1/0/1 | in output drop

Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: **497000** 預設閾值Th2中可見這些丟包。預設情況下,系統可以使用最大閾值作為丟棄閾值,即Drop-Th2。

3800#show platform qos queue stats gigabitEthernet 1/0/1

497000 0 0 1 0 0 0 0 0

然後,配置此服務策略以調整緩衝區:

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 gos queue config gigabitEthernet 2/0/1

3850#**show interfaces gigabitEthernet1/0/1 | include output drop** Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: **385064**

對於相同的流量突發,丟棄數從497000減少到385064。然而,仍有下降。之後,配置 qos queuesoftmax-multiplier 1200 global config 指令。

3850#**show interfaces gigabitEthernet1/0/1 | in output drop** Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: 0 **queue-0的Softmax最多可以達到10.000個緩衝區,因此丟棄數為零。**

註:這種情形是不可能的,因為其他介面也可以使用緩衝區,但這肯定有助於將資料包丟棄降 低到某個水準。

使用此命令可以增加介面可用的最大軟緩衝區,但是您也必須牢記,僅當沒有其它介面使用這些緩 衝區時,該軟緩衝區才可用。

摘要

1.建立更多隊列時,每個隊列獲得的緩衝區會更少。

2.可用緩衝區總數可透過以下方式增加: gos queue-softmax-multiplier 指令。

3.如果只定義1類預設值,為了調整緩衝區,所有流量都歸入單個隊列(包括控制資料包)中。請注 意,當所有流量都放在一個隊列中時,控制流量和資料流量之間沒有分類,在擁塞期間,控制流量 可能會被丟棄。因此,建議至少為控制流量建立其他1個類。CPU生成的控制資料包始終轉到第一 個優先順序隊列,即使類別對映中不匹配也是如此。如果沒有配置優先順序隊列,它將轉到介面的 第一個隊列,即queue-0。

4.在思科錯誤ID <u>CSCuu14019</u>之前,介面不會顯示「輸出丟棄」計數器。您必須執行 show platform gos queue stats 輸出以檢查丟包。

5.已提交增強請求Cisco錯誤ID <u>CSCuz8625</u>,以便我們在不使用任何服務策略的情況下配置softmax乘數。(已在3.6.6及更高版本中解決)

相關資訊

• <u>技術支援 - Cisco Systems</u>

關於此翻譯

思科已使用電腦和人工技術翻譯本文件,讓全世界的使用者能夠以自己的語言理解支援內容。請注 意,即使是最佳機器翻譯,也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準 確度概不負責,並建議一律查看原始英文文件(提供連結)。