

Quedas de saída do Nexus 3500 e QoS de buffer

Contents

[Introduction](#)

[Metodologia](#)

[Verifique se há quedas de saída](#)

[Determine se as quedas são unicast ou multicast](#)

[Determine qual buffer de saída é usado](#)

[Verifique o monitoramento de buffer ativo](#)

[Incremento Ativo de Contadores](#)

[Breve saída](#)

[Saída detalhada](#)

[Gerar um registro quando um limite é cruzado](#)

[IDs de bug notáveis da Cisco](#)

[Perguntas mais freqüentes](#)

[Apêndice - Informações do recurso](#)

[Gerenciamento de buffer](#)

[Programação](#)

[Receptor lento multicast](#)

[Monitoramento de buffer ativo](#)

[Implementação de hardware](#)

[Implementação de software](#)

Introduction

Este documento descreve os comandos usados para solucionar problemas do tipo de tráfego descartado na plataforma Nexus 3500 e no buffer de saída (OB) no qual esse tráfego é descartado.

Metodologia

1. [Verifique se há quedas de saída](#)
2. [Determine se as quedas são unicast ou multicast](#)
3. [Determine qual buffer de saída é usado](#)
4. [Verifique o monitoramento de buffer ativo](#)

Verifique se há quedas de saída

Verifique as estatísticas da interface física para determinar se o tráfego é descartado na direção de saída. Determine se o contador "output discard" na direção TX é incrementado e/ou diferente de zero.

```

Nexus3548# show interface Eth1/7
Ethernet1/7 is up
Dedicated Interface
Hardware: 100/1000/10000 Ethernet, address: a44c.116a.913c (bia a44c.116a.91ee)
Description: Unicast Only
Internet Address is 1.2.1.13/30
MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec
reliability 255/255, txload 35/255, rxload 1/255
Encapsulation ARPA
full-duplex, 1000 Mb/s, media type is 1G
Beacon is turned off
Input flow-control is off, output flow-control is off
Rate mode is dedicated
Switchport monitor is off
EtherType is 0x8100
Last link flapped 00:03:48
Last clearing of "show interface" counters 00:03:55
1 interface resets
30 seconds input rate 200 bits/sec, 0 packets/sec
30 seconds output rate 0 bits/sec, 0 packets/sec
Load-Interval #2: 5 minute (300 seconds)
  input rate 40 bps, 0 pps; output rate 139.46 Mbps, 136.16 Kpps
RX
  1 unicast packets  118 multicast packets  0 broadcast packets
  119 input packets  9830 bytes
  0 jumbo packets  0 storm suppression bytes
  0 runts  0 giants  0 CRC  0 no buffer
  0 input error  0 short frame  0 overrun  0 underrun  0 ignored
  0 watchdog  0 bad etype drop  0 bad proto drop  0 if down drop
  0 input with dribble  0 input discard
  0 Rx pause
TX
  23605277 unicast packets  0 multicast packets  0 broadcast packets
  23605277 output packets  3038908385 bytes
  0 jumbo packets
  0 output errors  0 collision  0 deferred  0 late collision
  0 lost carrier  0 no carrier  0 babble 11712542 output discard
  0 Tx pause

```

Determine se as quedas são unicast ou multicast

Quando for determinado que a interface descarta o tráfego, insira o comando **show queuing interface <x/y>** para descobrir se o tráfego descartado é multicast ou unicast. Nas versões anteriores a 6.0(2)A3(1), a saída é semelhante a:

```

Nexus3548# show queuing interface Eth1/7
Ethernet1/7 queuing information:
TX Queuing
  qos-group  sched-type  oper-bandwidth
    0          WRR          100

RX Queuing
  Multicast statistics:
    Mcast pkts dropped          : 0
  Unicast statistics:
    qos-group 0
    HW MTU: 1500 (1500 configured)
    drop-type: drop, xon: 0, xoff: 0
  Statistics:
    Ucast pkts dropped          : 11712542

```

Na versão 6.0(2)A3(1) e posterior, a saída é semelhante a:

```
Nexus3548# show queuing interface Eth1/7
Ethernet1/7 queuing information:
  qos-group sched-type oper-bandwidth
    0         WRR         100
Multicast statistics:
  Mcast pkts dropped          : 0
Unicast statistics:
  qos-group 0
  HW MTU: 1500 (1500 configured)
  drop-type: drop, xon: 0, xoff: 0
  Statistics:
Ucast pkts dropped          : 11712542
```

Note: Se o receptor lento de multicast estiver configurado para a porta, consulte para obter informações sobre o recurso, as quedas não serão rastreadas com o comando `show queuing interface Eth<x/y>` devido a uma limitação de hardware. Consulte o bug da Cisco ID [CSCuj21006](https://tools.cisco.com/bugcenter/bug/?bugID=CSCuj21006).

Determine qual buffer de saída é usado. No Nexus 3500, há três pools de buffer usados na direção de saída. A saída do comando `show hardware internal mtc-usr info port-mapping` fornece as informações de mapeamento.

```
Nexus3548# show hardware internal mtc-usr info port-mapping
```

OB Ports to Front Ports:

```
===== OB0 =====      ===== OB1 =====      ===== OB2 =====
45 47 21 23 09 11 33 35    17 19 05 07 41 43 29 31    13 15 37 39 25 27 01 03
46 48 22 24 10 12 34 36    18 20 06 08 42 44 30 32    14 16 38 40 26 28 02 04
```

Front Ports to OB Ports:

```
=OB2= =OB1= =OB0= =OB2=    =OB1= =OB0= =OB2= =OB1=    =OB0= =OB2= =OB1= =OB0=
12 14 04 06 08 10 00 02    00 02 04 06 08 10 12 14    12 14 04 06 08 10 00 02
13 15 05 07 09 11 01 03    01 03 05 07 09 11 13 15    13 15 05 07 09 11 01 03
```

Front port numbering (i.e. "01" here is e1/1):

```
=OB2= =OB1= =OB0= =OB2=    =OB1= =OB0= =OB2= =OB1=    =OB0= =OB2= =OB1= =OB0= 01 03 05 07 09 11
13 15    17 19 21 23 25 27 29 31    33 35 37 39 41 43 45 47 02 04 06 08 10 12 14 16    18 20 22
24 26 28 30 32    34 36 38 40 42 44 46 48
```

Note: Text in Red font is not CLI output, it's purely to help those reading the document faster match the actual front port instead of having to manually count up.

A primeira parte dos resultados indica que o pool OB 0 é usado por portas frontais como 45, 46, 47, 48 e assim por diante, e OB1 é usado pelas portas frontais 17, 18 e assim por diante. A segunda parte dos resultados indica que Eth1/1 está mapeado para a porta OB2 12, Eth1/2 está mapeado para a porta OB2 13 e assim por diante. A porta em discussão, Eth1/7, é mapeada para OB1. Consulte a seção [Gerenciamento de Buffer](#) neste documento para obter mais informações. Verifique o monitoramento de buffer ativo. Consulte o whitepaper [Cisco Nexus 3548 Ative Buffer Monitoring](#) e a seção neste documento para obter mais informações sobre esse recurso. **Incremento Ativo de Contadores** Se a saída descartar o incremento ativo, habilite o Ative Buffer Monitoring (ABM) com esse comando. Observe que o comando permite monitorar unicast ou multicast, mas não ambos. Além disso, permite configurar o intervalo de amostragem e os valores de limite.

```
hardware profile buffer monitor [unicast|multicast] {[sampling
```

Breve saída Quando o ABM estiver ativado, você poderá exibir os resultados com esse comando.

```
Nexus3500# show hardware profile buffer monitor interface e1/7 brief
```

```
Brief CLI issued at: 09/30/2013 19:43:50
```

```
Maximum buffer utilization detected
1sec      5sec      60sec     5min      1hr
-----
Ethernet1/7 5376KB  5376KB  5376KB    N/A      N/A
```

Esses resultados indicam que 5,376 MB de 6 MB do buffer OB1 foram usados pelo tráfego unicast que deixou Eth1/7 nos últimos 60 segundos. Saída detalhada

```
Nexus3500# show hardware profile buffer monitor interface Eth1/7 detail
```

```
Detail CLI issued at: 09/30/2013 19:47:01
```

Legend -

384KB - between 1 and 384KB of shared buffer consumed by port

768KB - between 385 and 768KB of shared buffer consumed by port

307us - estimated max time to drain the buffer at 10Gbps

Active Buffer Monitoring for Ethernet1/7 is: Active

```
KBytes      384  768 1152 1536 1920 2304 2688 3072 3456 3840 4224 4608 4992 5376 5760
6144
us @ 10Gbps  307  614  921 1228 1535 1842 2149 2456 2763 3070 3377 3684 3991 4298 4605
4912
```

```
-----
09/30/2013 19:47:01    0    0    0    0    0    0    0    0    0    0    0    0    0    0    250
0    0
09/30/2013 19:47:00    0    0    0    0    0    0    0    0    0    0    0    0    0    0    252
0    0
09/30/2013 19:46:59    0    0    0    0    0    0    0    0    0    0    0    0    0    0    253
0    0
09/30/2013 19:46:58    0    0    0    0    0    0    0    0    0    0    0    0    0    0    250
0    0
09/30/2013 19:46:57    0    0    0    0    0    0    0    0    0    0    0    0    0    0    250
0    0
09/30/2013 19:46:56    0    0    0    0    0    0    0    0    0    0    0    0    0    0    250
0    0
09/30/2013 19:46:55    0    0    0    0    0    0    0    0    0    0    0    0    0    0    251
0    0
09/30/2013 19:46:54    0    0    0    0    0    0    0    0    0    0    0    0    0    0    251
0    0
09/30/2013 19:46:53    0    0    0    0    0    0    0    0    0    0    0    0    0    0    250
0    0
09/30/2013 19:46:52    0    0    0    0    0    0    0    0    0    0    0    0    0    0    253
0    0
09/30/2013 19:46:51    0    0    0    0    0    0    0    0    0    0    0    0    0    0    249
0    0
...
```

As informações em cada linha são registradas em um segundo intervalo. Cada coluna representa o uso do buffer. Conforme mencionado nos resultados do comando, se houver um valor diferente de zero reportado para a coluna "384", significa que o uso do buffer estava entre 0 e 384 KBytes quando o ABM pesquisou o uso do OB. O número diferente de zero é o número de vezes que o uso foi relatado. Esses resultados indicam que o OB1 alcançou uma média de 5,376 MB de uso entre 249 e 253 vezes por segundo nos últimos 10 segundos para Eth1/7. Leva 4298 microssegundos (us) para limpar o buffer desse tráfego. Gerar um registro quando um limite é cruzado Se o contador de queda e o uso do buffer incrementarem periodicamente, será possível definir um limite e gerar uma mensagem de log quando o limite for ultrapassado.

```
logging level mtc-usd 5
```

```
hardware profile buffer monitor unicast sampling 10 threshold 4608
```

O comando é definido para monitorar o tráfego unicast em um intervalo de 10 nanossegundos e quando ele vai além de 75% do buffer, ele gera um log. Você também pode criar um agendador para coletar estatísticas ABM e saída do contador de interface a cada hora e anexá-lo aos arquivos bootflash. Este exemplo é para tráfego multicast:

```
hardware profile buffer monitor multicast
```

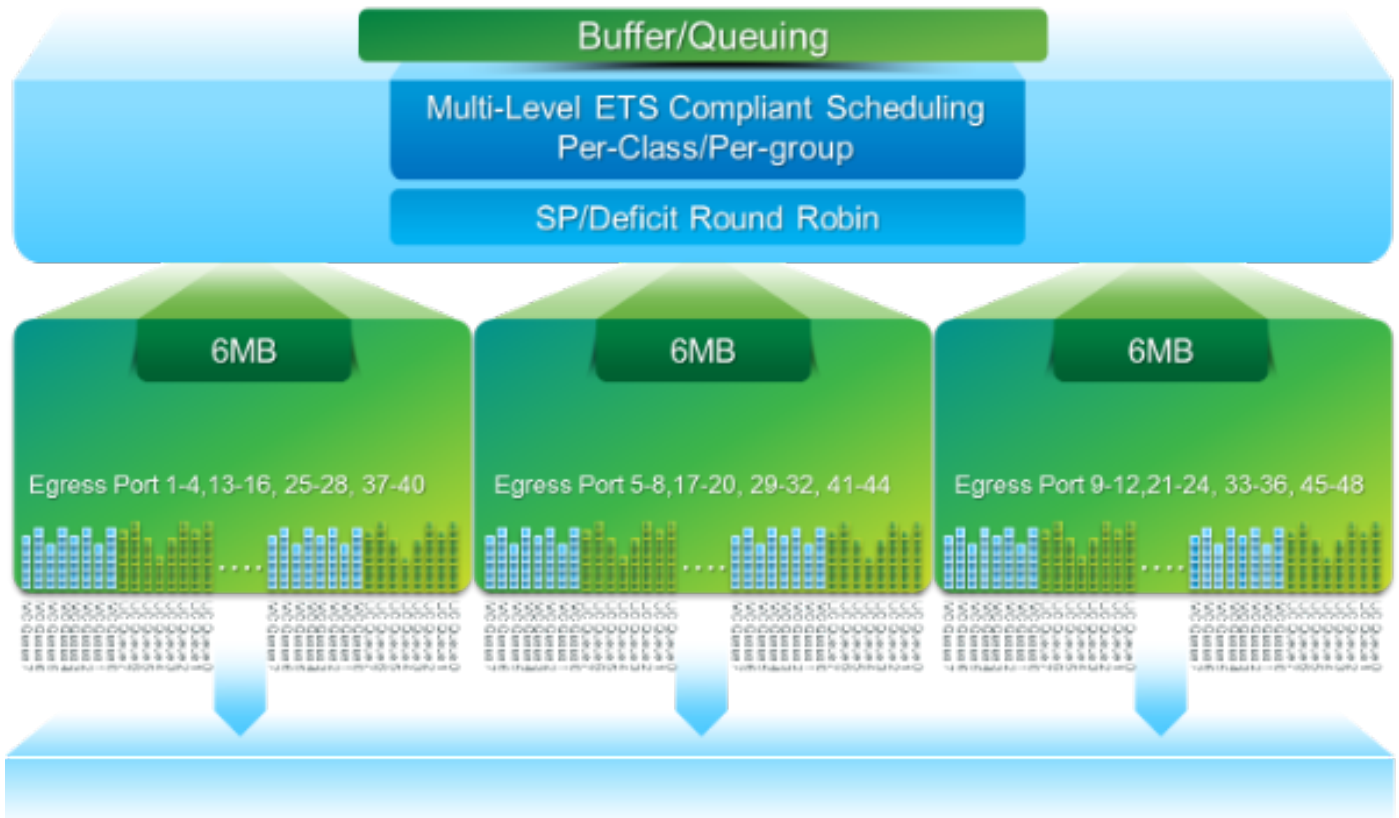
```
feature scheduler
scheduler job name ABM
show hardware profile buffer monitor detail >> ABMDetail.txt
show clock >> ABMBrief.txt
show hardware profile buffer monitor brief >> ABMBrief.txt
show clock >> InterfaceCounters.txt
show interface counters errors >> InterfaceCounters.txt
scheduler schedule name ABM
time start now repeat 1:0
job name ABM
```

IDs de bug notáveis da Cisco

- ID de bug da Cisco [CSCum21350](#): As oscilações de porta rápidas fazem com que todas as portas no mesmo buffer de QoS descartem todo o tráfego de transmissão/multicast de TX. Isso é corrigido na versão 6.0(2)A1(1d) e posterior.
- ID de bug da Cisco [CSCuq96923](#): O bloco de buffer multicast está travado, o que resulta em quedas de multicast/broadcast de saída. Esta questão ainda está a ser investigada.
- ID de bug da Cisco [CSCva20344](#): Bloco/travamento do buffer do Nexus 3500 - sem transmissão ou multicast TX. Problema inreproduzível, potencialmente corrigido nas versões 6.0(2)U6(7), 6.0(2)A6(8) e 6.0(2)A8(3).
- ID de bug da Cisco [CSCvi93997](#): bloco de buffer de saída dos switches Cisco Nexus 3500 travado. Isso é corrigido nas versões 7.0(3)I7(8) e 9.3(3).

Perguntas mais frequentes O ABM afeta o desempenho ou a latência? Não, esse recurso não afeta a latência ou o desempenho do dispositivo. Qual é o impacto do menor intervalo de pesquisa de hardware ABM? Por padrão, o intervalo de sondagem de hardware é de 4 milissegundos. Você pode configurar esse valor como 10 nanossegundos. Não há impacto no desempenho ou na latência devido ao menor intervalo de pesquisa de hardware. A pesquisa de hardware padrão de 4 milissegundos é selecionada para garantir que você não sobrecarregue os contadores de histograma antes das pesquisas de software a cada segundo. Se você reduzir o intervalo de pesquisa de hardware, ele poderá saturar os contadores de hardware em 255 amostras. O dispositivo não pode lidar com uma pesquisa de software inferior a um segundo, para corresponder à pesquisa de hardware mais baixa devido a restrições de CPU e memória. O whitepaper tem o exemplo do intervalo de pesquisa de hardware mais baixo e seu caso de

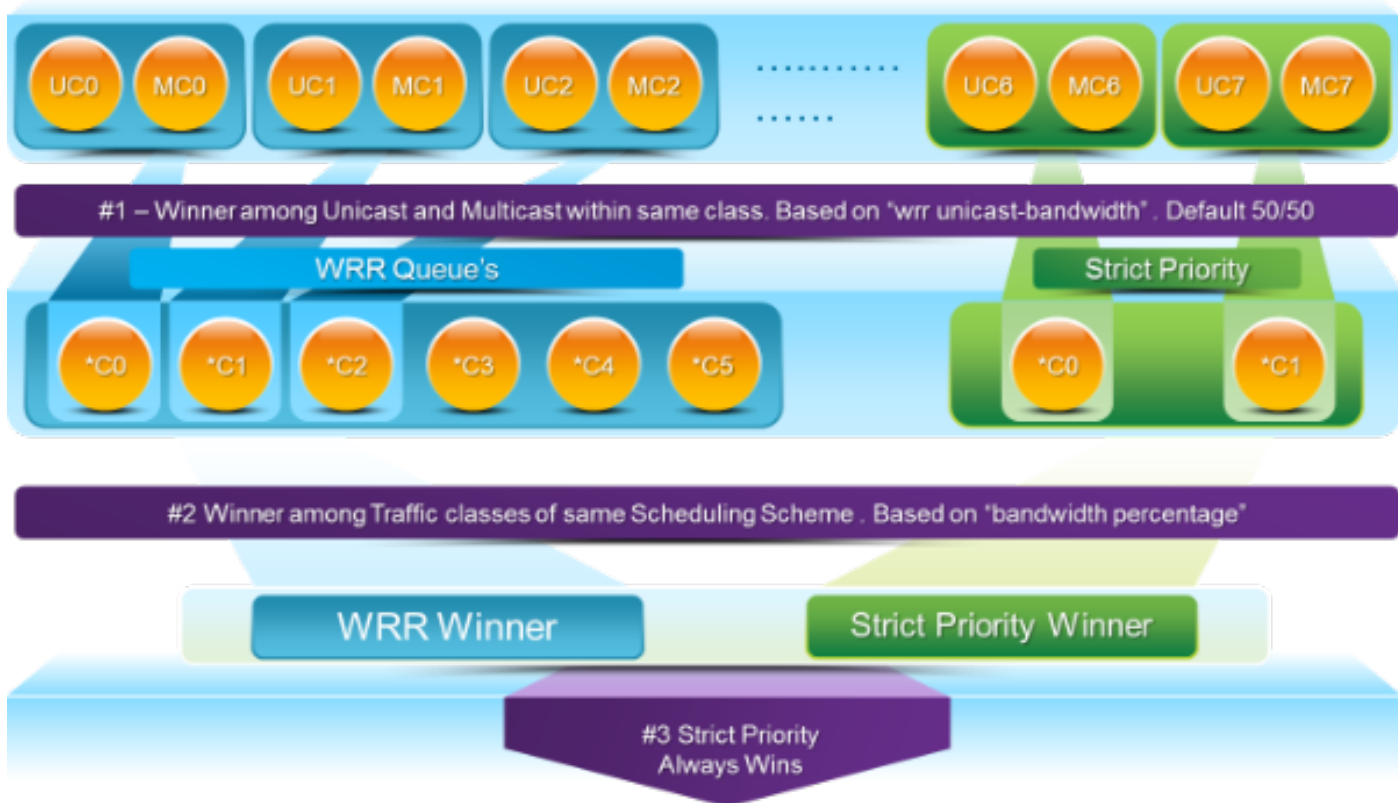
uso. **Apêndice - Informações do recurso** Gerenciamento de buffer



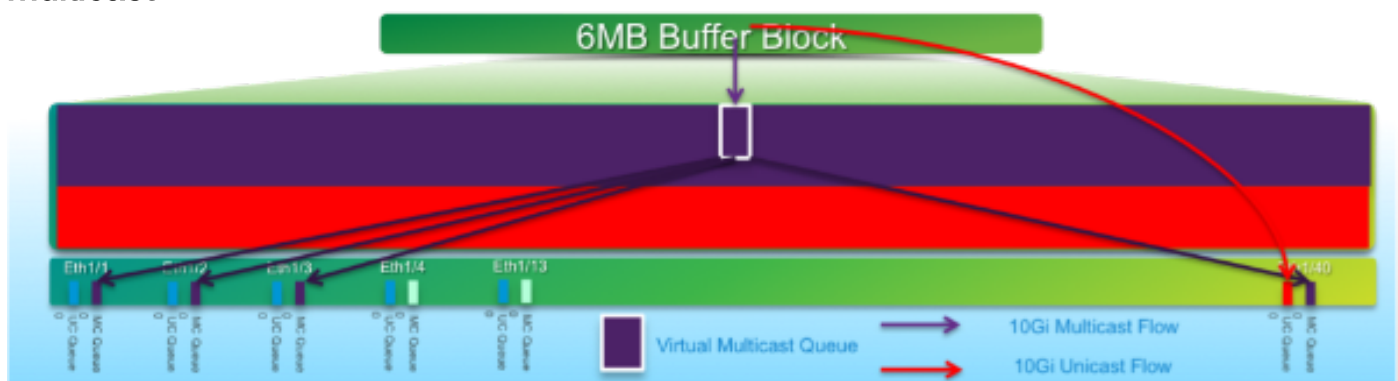
- Buffer de pacote de 18 MB compartilhado por três blocos OB: ~4 MB reservados: Tamanho com base na MTU (Maximum Transmission Unit, Unidade máxima de transmissão) configurada (soma por porta de $2 \times \text{MTU Size} \times \# \text{ de grupos de QoS habilitados}$) ~14 MB compartilhados: Resto do buffer total ~767 KB de OB: 0 para pacotes destinados à CPU
- 6 MB para cada OB é compartilhado por um conjunto de 16 portas (comando `show hardware internal mtc-usd info port-mapping`)

Programação Programação em três camadas:

- Unicast e multicast
- Classes de tráfego do mesmo esquema de programação
- Classes de tráfego no esquema



Receptor lento multicast



Neste diagrama:

- Congestionamento sustentado é introduzido em 1 G Eth1/40.
- Outros receptores multicast (Eth1/1 - 3) no bloco de buffer são afetados devido ao comportamento de agendamento multicast. *Os receptores em outros blocos de buffer não são afetados.*
- O "receptor lento multicast" pode ser aplicado a e1/40 para evitar perda de tráfego em portas não congestionadas.
- O "receptor lento multicast" esgota o multicast a uma taxa de 10 G em Eth1/40. *Ainda é esperado que ocorram quedas na porta congestionada.*
- Configurado com o comando multicast do perfil de hardware, low-receptor port <x>.

Monitoramento de buffer ativo Consulte o whitepaper [Cisco Nexus 3548 Active Buffer Monitoring](#) para obter uma visão geral desse recurso. Implementação de hardware

- O ASIC tem 18 buckets e cada bucket corresponde a um intervalo de utilização de buffer (por exemplo, 0-384KB, 385-768KB e assim por diante).
- O ASIC pesquisa a utilização do buffer para todas as portas a cada 4 milissegundos (padrão). Esse intervalo de sondagem ASIC pode ser configurado em até 10 nanossegundos.

- Com base na utilização do buffer para cada intervalo de pesquisa de hardware, o contador de bucket para o intervalo correspondente é incrementado. Ou seja, se a porta 25 consumir 500 KB de buffer, o contador de bucket 2 (385-768 KB) é incrementado.
- Esse contador de utilização de buffer é mantido para cada interface no formato de histograma.
- Cada bucket é representado por 8 bits, de modo que o contador atinge o valor máximo em 255 e é redefinido quando o software lê os dados.

Implementação de software

- A cada segundo, o software pesquisa o ASIC para baixar e limpar todos os contadores de histograma.
- Esses contadores de histograma são mantidos na memória por 60 minutos com granularidade de um segundo.
- O software também garante que copie o histograma do buffer no flash de inicialização a cada hora, que pode ser copiado para o analisador para análise posterior.
- Efetivamente, isso mantém dois horas de dados de histograma de buffer para todas as portas, a última hora na memória e a segunda hora no flash de inicialização.