

Nexus 3500 - Ausgabeverluste und Buffer-QoS

Inhalt

[Einführung](#)

[Methodik](#)

[Auf Ausgabeverluste prüfen](#)

[Stellen Sie fest, ob die Drops Unicast oder Multicast sind.](#)

[Bestimmen Sie, welcher Ausgabepuffer verwendet wird.](#)

[Überprüfen Sie die aktive Puffer-Überwachung.](#)

[Zähler erhöhen sich aktiv](#)

[Kurze Ausgabe](#)

[Detaillierte Ausgabe](#)

[Erstellen eines Protokolls beim Überschreiten eines Schwellenwerts](#)

[Wichtige Cisco Bug-IDs](#)

[Häufig gestellte Fragen](#)

[Anhang - Informationen zu Funktionen](#)

[Puffer-Management](#)

[Planung](#)

[Multicast-Langsamempfänger](#)

[Aktive Puffer-Überwachung](#)

[Hardware-Implementierung](#)

[Softwareimplementierung](#)

Einführung

In diesem Dokument werden Befehle beschrieben, die zur Fehlerbehebung bei der Art des auf der Nexus 3500-Plattform verworfenen Datenverkehrs und dem Ausgabepuffer (OB) verwendet werden, in dem dieser Datenverkehr verworfen wird.

Methodik

1. [Auf Ausgabeverluste prüfen](#)
2. [Stellen Sie fest, ob die Drops Unicast oder Multicast sind.](#)
3. [Bestimmen Sie, welcher Ausgabepuffer verwendet wird.](#)
4. [Überprüfen Sie die aktive Puffer-Überwachung.](#)

Auf Ausgabeverluste prüfen

Überprüfen Sie die Statistiken zur physischen Schnittstelle, um festzustellen, ob der Datenverkehr in die Ausgangsrichtung verworfen wird. Stellen Sie fest, ob der Zähler für "Output Discard" in TX-Richtung inkrementiert und/oder nicht 0 ist.

```

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

```

Stellen Sie fest, ob die Drops Unicast oder Multicast sind.

Wenn festgestellt wurde, dass die Schnittstelle den Datenverkehr verwirft, geben Sie den Befehl **show queuing interface <x/y>** ein, um herauszufinden, ob der Datenverkehr als Multicast oder Unicast verworfen wird. In Versionen vor 6.0(2)A3(1) sieht die Ausgabe wie folgt aus:

```

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

```

In Version 6.0(2)A3(1) und höher sieht die Ausgabe wie folgt aus:

```

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

```

Hinweis: Wenn der Multicast Slow Receiver für den Port konfiguriert ist, finden Sie Informationen zu den Funktionen, und Verwerfungen werden aufgrund einer Hardware-Einschränkung nicht mit dem Befehl `show queuing interface Eth<x/y>` nachverfolgt. Siehe Cisco Bug ID [CSCuj21006](#).

Bestimmen Sie, welcher Ausgabepuffer verwendet wird. Auf dem Nexus 3500 werden drei Pufferpools in Ausgangsrichtung verwendet. Die Ausgabe des Befehls `show hardware internal mtc-usd info port-mapping` liefert die Zuordnungsinformationen.

```

Nexus3548# show hardware internal mtc-usd 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.

Der erste Teil der Ergebnisse zeigt an, dass der OB-Pool 0 von den Front-Ports wie 45, 46, 47, 48 usw. verwendet wird, und OB1 von den Front-Ports 17, 18 usw. Der zweite Teil der Ergebnisse zeigt an, dass Eth1/1 dem OB2-Port 12 zugeordnet ist, Eth1/2 dem OB2-Port 13 usw. zugeordnet ist. Der diskutierte Port Eth1/7 ist OB1 zugeordnet. Weitere Informationen finden Sie im Abschnitt [Buffer Management](#) in diesem Dokument. Überprüfen Sie die aktive Puffer-

Überwachung. Weitere Informationen zu dieser Funktion finden Sie im Whitepaper [Cisco Nexus 3548 zur aktiven Puffer-Überwachung](#) und im Abschnitt in diesem Dokument. Zähler erhöhen sich aktiv. Wenn die Ausgabe aktiv inkrementiert wird, aktivieren Sie mit diesem Befehl die Option Active Buffer Monitoring (ABM). Beachten Sie, dass Sie mit dem Befehl Unicast oder Multicast überwachen können, aber nicht beides. Außerdem können Sie das Samplingintervall und die Schwellenwerte konfigurieren.

```

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

```

Kurze Ausgabe Sobald der ABM aktiviert ist, können Sie die Ergebnisse mit diesem Befehl anzeigen.

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

Diese Ergebnisse zeigen, dass 5,376 MB von 6 MB des OB1-Puffers von Unicast-Datenverkehr verwendet wurden, der Eth1/7 während der letzten 60 Sekunden zurückgelassen hat. Detaillierte Ausgabe

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															

	384	768	1152	1536	1920	2304	2688	3072	3456	3840	4224	4608	4992	5376	5760
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															

Die Informationen in jeder Zeile werden in einem zweiten Intervall protokolliert. Jede Spalte stellt die Puffernutzung dar. Wie in den Befehlsergebnissen erwähnt, bedeutet dies, dass die Puffernutzung zwischen 0 und 384 KByte betrug, wenn der ABM die OB-Nutzung abfragt, wenn für die Spalte "384" ein Nicht-Nullwert gemeldet wurde. Die Nicht-Null-Zahl ist die Anzahl der Fälle, in denen die Nutzung gemeldet wurde. Diese Ergebnisse zeigen, dass OB1 in den letzten 10 Sekunden für Eth1/7 zwischen dem 249- und dem 253-fachen der Nutzung pro Sekunde durchschnittlich 5,376 MB betrug. Es dauert 4298 Mikrosekunden (us), um den Puffer für diesen Datenverkehr zu löschen. Erstellen eines Protokolls beim Überschreiten eines Schwellenwerts Wenn der Drop-Zähler und die Puffernutzung in regelmäßigen Abständen erhöht werden, kann beim Überschreiten des Schwellenwerts ein Grenzwert festgelegt und eine

Protokollmeldung generiert werden.

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

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

Der Befehl ist so eingestellt, dass Unicast-Datenverkehr in einem Intervall von 10 Nanosekunden überwacht wird. Wenn dieser Wert über 75 % des Puffers liegt, wird ein Protokoll generiert. Sie können auch einen Scheduler erstellen, um stündlich ABM-Statistiken und Schnittstellen-Zählerausgabe zu erfassen und an Bootflash-Dateien anzuhängen. Dieses Beispiel bezieht sich auf Multicast-Datenverkehr:

```
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
```

Wichtige Cisco Bug-IDs

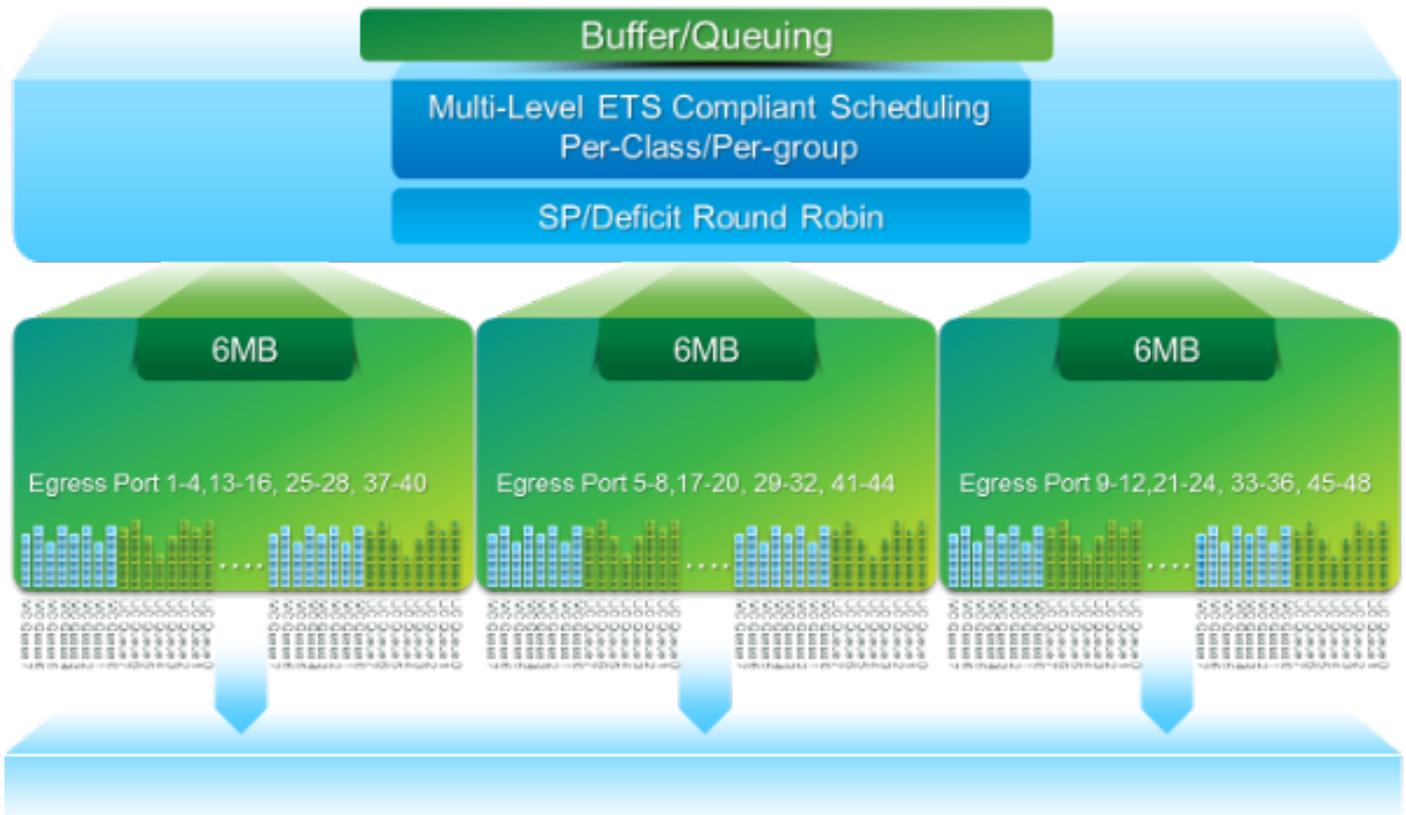
- Cisco Bug-ID [CSCum21350](#): Bei schnellen Port-Flaps werden alle Ports im gleichen QoS-Puffer deaktiviert, um den gesamten TX-Multicast-/Broadcast-Datenverkehr zu verwerfen. Dies wurde in Version 6.0(2)A1(1d) und höher behoben.
- Cisco Bug-ID [CSCuq96923](#): Der Multicast-Pufferblock ist blockiert, was Ausgangs-Multicast-/Broadcast-Verwerfungen zur Folge hat. Dieses Problem wird noch untersucht.
- Cisco Bug-ID [CSCva20344](#): Nexus 3500-Puffer-Block/Lockup - kein TX-Multicast oder Broadcast. Unreproduzierbares Problem, möglicherweise behoben in den Versionen 6.0(2)U6(7), 6.0(2)A6(8) und 6.0(2)A8(3).
- Cisco Bug ID [CSCvi93997](#): Ausgabepufferblock der Cisco Nexus Switches der Serie 3500 blockiert. Dies ist in den Versionen 7.0(3)I7(8) und 9.3(3) behoben.

Häufig gestellte Fragen Wirkt sich ABM auf Leistung oder Latenz aus? Nein, diese Funktion beeinträchtigt weder die Latenz noch die Leistung des Geräts. Welche Auswirkungen hat das niedrigere ABM-Hardware-Polling-Intervall? Das Hardware-Polling-Intervall beträgt standardmäßig 4 Millisekunden. Sie können diesen Wert bis zu 10 Nanosekunden einstellen. Aufgrund des niedrigeren Hardware-Polling-Intervalls treten keine Auswirkungen auf die Leistung oder Latenz auf. Das standardmäßige Hardware-Polling von 4 Millisekunden wird ausgewählt, um sicherzustellen, dass die Histogramm-Zähler nicht alle eine Sekunde vor den Softwareabfragen überlaufen. Wenn Sie das Hardware-Polling-Intervall reduzieren, können die Hardwareindikatoren mit 255 Beispielen ausgelastet werden. Das Gerät kann ein Software-Polling von weniger als einer Sekunde nicht verarbeiten, um die niedrigere Hardware-Polling aufgrund von CPU- und Speicherbeschränkungen abzugleichen. Das Whitepaper zeigt ein Beispiel für das niedrigere

Hardware-Polling-Intervall und dessen Anwendungsfall.

Anhang - Informationen zu Funktionen

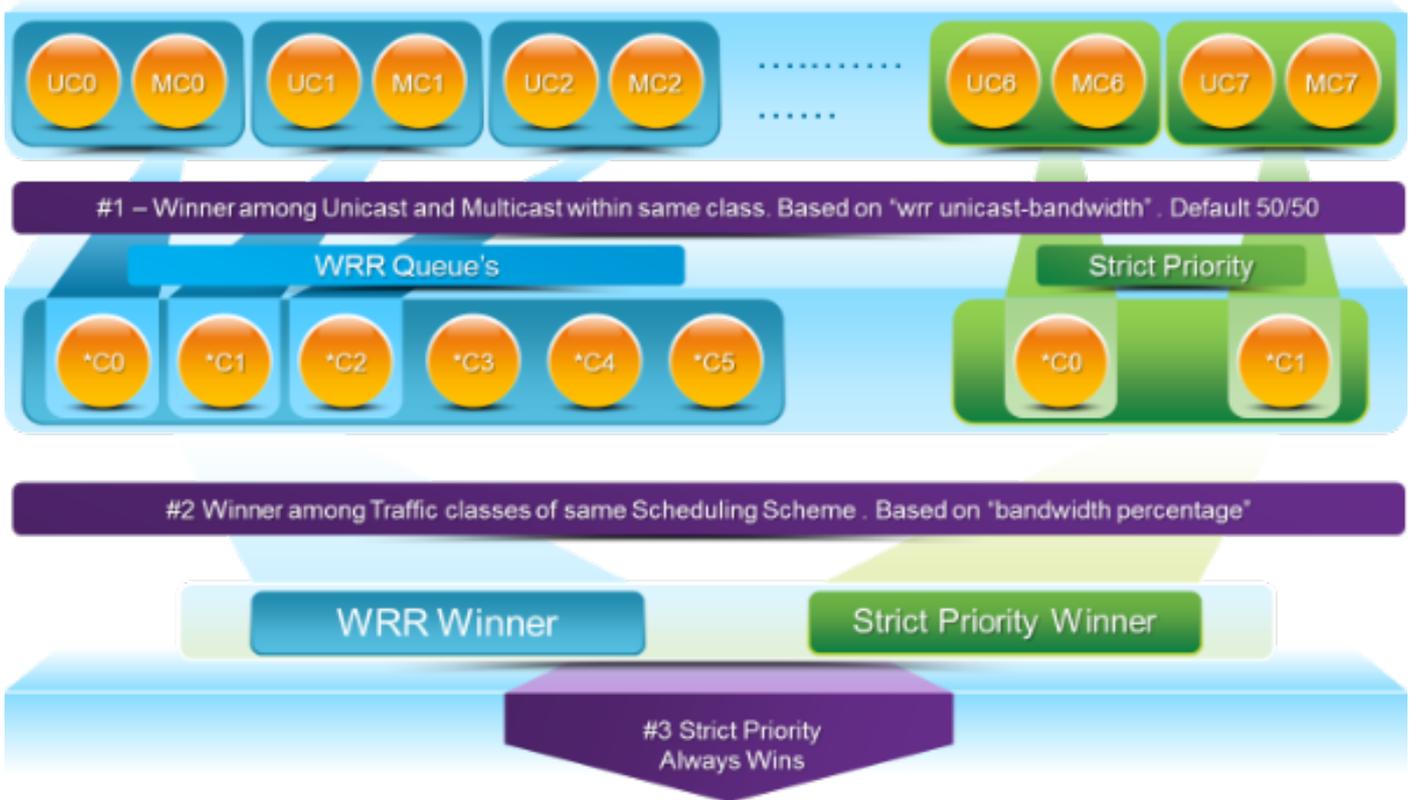
Puffer-Management



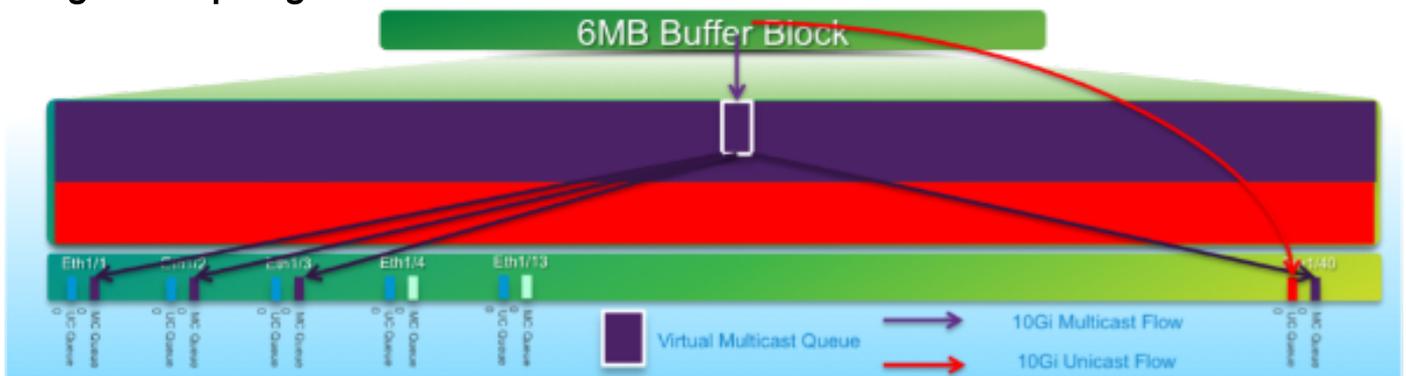
- 18 MB Paketpuffer, der von drei OB-Blöcken gemeinsam genutzt wird: ~4 MB reserviert: Größe basierend auf konfigurierter Maximum Transmission Unit (MTU) (Pro Port-Summe von 2 x MTU-Größe x Anzahl aktivierter QoS-Gruppen)~14 MB gemeinsam genutzt: Rest des Gesamtpuffers~767 KB OB: 0 für CPU-Zielpakete
- 6 MB pro OB werden von einem Satz von 16 Ports gemeinsam genutzt (zeigen den hardwareinternen mtc-usb info port-mapping-Befehl an)

PlanungDreistufige Planung:

- Unicast und Multicast
- Datenverkehrsklassen desselben Scheduling-Schemas
- Datenverkehrsklassen im gesamten Schema



Multicast-Langsamempfänger



In diesem Diagramm:

- Bei 1 G Eth1/40 wird eine anhaltende Überlastung festgestellt.
- Andere Multicast-Empfänger (Eth1/1-3) im Pufferblock sind aufgrund des Multicast-Planungsverhaltens betroffen. *Empfänger in anderen Pufferblöcken bleiben davon unberührt.*
- "Multicast-Slow-Receiver" kann auf e1/40 angewendet werden, um Datenverkehrsverluste an nicht überlasteten Ports zu vermeiden.
- Der "Multicast Slow-Receiver" leitet Multicast bei Eth1/40 mit einer Geschwindigkeit von 10 G ab. *Es wird weiterhin erwartet, dass es auf dem überlasteten Port zu Verlusten kommt.*
- Konfiguriert mit dem Befehl `hardware profile multicast slow-Receiver port <x>`.

Aktive Puffer-Überwachung Eine Übersicht dieser Funktion finden Sie im Whitepaper [Cisco Nexus 3548 zur aktiven Puffer-Überwachung](#). Hardware-Implementierung

- ASIC hat 18 Eimer und jeder Eimer entspricht einem Bereich der Puffernutzung (z. B. 0-384 KB, 385-768 KB usw.).
- Der ASIC fragt die Puffernutzung für alle Ports alle 4 Millisekunden ab (Standard). Dieses ASIC-Abfrageintervall kann bis zu 10 Nanosekunden konfiguriert werden.
- Basierend auf der Puffernutzung für jedes Hardware-Polling-Intervall wird der Ecket-Zähler für den entsprechenden Bereich erhöht. Das heißt, wenn Port 25 500 KB Puffer verbraucht, wird

der Zähler der Bucket #2 (385-768 KB) erhöht.

- Dieser Zähler für die Puffernutzung wird für jede Schnittstelle im Histogramm-Format beibehalten.
- Jeder Eimer wird mit 8 Bit dargestellt, sodass der Zähler mit 255 abgebildet ist und zurückgesetzt wird, sobald die Software die Daten liest.

Softwareimplementierung

- Jede Sekunde fragt die Software ASIC ab, um alle Histogramm-Zähler herunterzuladen und zu löschen.
- Diese Histogramm-Zähler werden 60 Minuten lang mit einer Genauigkeit von einer Sekunde im Speicher aufbewahrt.
- Die Software stellt außerdem sicher, dass das Pufferhistogramm stündlich in den Bootflash kopiert wird, der zur weiteren Analyse in den Analyzer kopiert werden kann.
- Auf diese Weise werden für alle Ports Pufferhistogramm-Daten im Wert von zwei Stunden gespeichert, die neueste Stunde im Speicher und die zweite Stunde im Bootflash.