

Umgang mit Microbursts in ASR 920, Unterschiede zwischen Bytes und Prozent mit Warteschlangenlimit

Inhalt

[Einleitung](#)

[Voraussetzungen](#)

[Anforderungen](#)

[Verwendete Komponenten](#)

[Hintergrundinformationen](#)

[Warteschlangenlimit konfigurieren](#)

[Unterschiede zu RSP3 und NCS 520](#)

[Warteschlangenlimit \(QL\), Lab-Test, Unterschiede zwischen Byte und Prozent](#)

[Testumfang und -parameter](#)

[Testverfahren](#)

[Beispiel](#)

[Testergebnisse](#)

[QL-Prozent vs. Bytes mit 64-Byte-Paketen](#)

[QL Prozent vs. Bytes mit 200 Bytes Packets](#)

[QL Prozent vs. Bytes mit 300 Bytes Packets](#)

[QL Prozent vs. Bytes mit 518 Bytes Packets](#)

[QL Prozent vs. Bytes mit 800 Bytes Packets](#)

[QL Prozent vs. Bytes mit 1024 Bytes Packets](#)

[QL Prozent vs. Bytes mit 1400 Bytes Packets](#)

[Beziehung zwischen Tropfen und Paketgröße](#)

[Differenz zwischen äquivalenten Bytes und Prozentwerten](#)

[Schlussfolgerungen](#)

[Wie man die Prozentwerte in realen Verkehrsszenarien annimmt](#)

[Überprüfen der Verwendung des Microburst-Puffers](#)

Einleitung

In diesem Dokument wird die Behandlung von Microbursts auf ASR 920-Routern beschrieben, die häufig die Ursache für das Verwerfen von Schnittstellenausgabepaketten sind. Im Detail wird mit dem Befehl `queue-limit` der Unterschied zwischen der Byteauslastung und dem Prozentwert dargestellt.

Voraussetzungen

Anforderungen

Cisco empfiehlt, dass Sie über Kenntnisse in folgenden Bereichen verfügen:

Router der Serie ASR 920

QoS-Richtlinien

Verwendete Komponenten

Die Informationen in diesem Dokument basieren auf einem ASR 920-Router, auf dem die Softwareversion Cisco IOS-XE 16.9.6 ausgeführt wird.

IXIA wird als Datenverkehrsgenerator für den Labortest verwendet.

Die Informationen in diesem Dokument beziehen sich auf Geräte in einer speziell eingerichteten Testumgebung. Alle Geräte, die in diesem Dokument benutzt wurden, begannen mit einer gelöschten (Nichterfüllungs) Konfiguration. Wenn Ihr Netzwerk in Betrieb ist, stellen Sie sicher, dass Sie die möglichen Auswirkungen aller Befehle verstehen.

Hintergrundinformationen

Microbursts werden als kleine Spitzen im Netzwerkverkehr bezeichnet. In Zugangsnetzwerken von Service Providern beispielsweise tritt dies häufig in Szenarien mit Geschwindigkeitsunterschieden auf, bei denen ein Datenverkehrsfluss von einer Hochgeschwindigkeits-Schnittstelle wie 10 Gigabit Ethernet (GE) in den Router gelangt und über eine Niedriggeschwindigkeits-Schnittstelle wie 1 GE ausgeht.

Das häufigste Problem, das durch Mikrobursts bei ASR 920- Routern verursacht wird, sind verworfene Schnittstellenpakete in Ausgaberrichtung. Dies tritt auf, wenn an der Eingangsschnittstelle für einen sehr kurzen Zeitraum (in der Größenordnung von Millisekunden) Bursts auftreten, die höher sind als die Rate der Ausgangsschnittstelle. Während dieser Zeit müssen die Pakete gepuffert werden.

Auf ASR 920/RSP2-Plattformen beträgt der Standardpuffer, der für Warteschlangen auf jeder 1-GE-Schnittstelle reserviert ist, 48 KB, während für die Warteschlangen auf jeder 10-GE-Schnittstelle 120 KB gilt.

Darüber hinaus steht ein gemeinsamer Puffer mit 11,75 MB zur Verfügung. Wenn der Puffer der Standardschnittstelle nicht ausreicht, um einen Paketerfolg zu ermöglichen, kann der freigegebene Puffer verwendet werden.

Um die Verwendung des freigegebenen Puffers für eine Schnittstelle zu aktivieren, muss unter dieser Schnittstelle eine QoS-Richtlinie konfiguriert werden, die den Parameter für das Warteschlangenlimit definiert.

Warteschlangenlimit konfigurieren

Dies ist ein Konfigurationsbeispiel für das Warteschlangenlimit der Schnittstelle GE 0/0/1:

```
class-map match-all DUMMY
match qos-group 99 <--- it can be any unused group
```

```
policy-map QUEUE-LIMIT
class DUMMY
class class-default
queue-limit percent 5
```

```
interface GigabitEthernet 0/0/1
service-policy output QUEUE-LIMIT
```

Diese Richtlinie gleicht keinen Datenverkehr ab und hat daher auch keine Auswirkungen auf diesen Datenverkehr. Diese Richtlinie erhöht lediglich den Warteschlangenpuffer der Schnittstelle GE 0/0/1.

Der Befehl **match qos-group 99** ist erforderlich, da die Konfiguration von Warteschlangenlimit in einer Nicht-Endknoten-Klasse nicht unterstützt wird. Daher müssen Sie eine fiktive Klassenzuordnung als übergeordnete Klasse verwenden, um den Befehl `queue-limit` unter dem Endknoten `class-default` zu konfigurieren:

```
ASR-920-1(config)#policy-map QUEUE-LIMIT
ASR-920-1(config-pmap)#class class-default
ASR-920-1(config-pmap-c)# queue-limit percent 5
QoS: queue-limit command not supported in non-leaf classes
queue-limit: platform params check fail
```

Die Warteschlangenbegrenzung kann auf verschiedene Weise festgelegt werden:

```
(config-pmap-c)#queue-limit ?
<1-2097152> in bytes, <1-1677721> in us, <1-8192000> in packets by default
percent % of threshol
```

Parameter und Wert müssen entsprechend Ihren Netzwerkanforderungen ausgewählt werden. Um diese Parameter zu verstehen und wie sich eine Änderung der Werte auf die Puffernutzung auswirkt, können Sie sich den im nächsten Abschnitt vorgestellten Labortest ansehen.

Anmerkung: Es wird nicht unterstützt, eine QoS-Richtlinie unter Port-Channel-Schnittstellen zu konfigurieren, sondern nur unter den physischen Schnittstellen des Port-Channels.

```
ASR-920-1(config)#interface port-channel 2
ASR-920-1(config-if)#service-policy output qos-tac
QoS: Configuration failed. Policy-map with Queueing actions not supported on EC main-
interface/EFPP
QoS: Configuration errors for policymap qos-tac
```

Unterschiede zu RSP3 und NCS 520

[Cisco RSP3 Module QoS-Funktionen:](#)

- Das RSP3-Modul verfügt über 4 GB externe Paketpuffer pro NPU.
- RSP3-Modul unterstützt 48000 Warteschlangen
- Standardmäßig unterstützt das RSP3-Modul bis zu 1 MB Warteschlangenlimit pro Warteschlange
- Der prozentuale Warteschlangengrenzwert wird von 1 GB der Gesamtpuffer als ausreichend angesehen.

Für Router mit RSP3-Supervisoren und NCS 520 ist die Anzahl der konfigurierbaren QoS-Gruppen auf 0-7 beschränkt:

```
ASR-903-1(config)#class-map match-all qos-tac
ASR-903-1(config-cmap)#match qos-group ?
```

<0-7> Qos Group value

Im NCS 520 gibt es standardmäßig einen 2 MB großen Warteschlangenpuffer, der von allen Schnittstellen gemeinsam genutzt wird. Auf einen externen 2 GB großen Puffer kann zugegriffen werden, wenn eine Richtlinienzuordnung mit Warteschlangenlimit konfiguriert wurde. Es gibt auch einen Unterschied in den Byte- und us-Parametern für das Warteschlangenlimit:

```
ASR-520-1(config-pmap-c)# queue-limit ?  
<1-8192000> in bytes, <1-40000> in us, <1-8192000> in packets by default  
percent % of threshold
```

Warteschlangenlimit (QL), Lab-Test, Unterschiede zwischen Byte und Prozent

Testumfang und -parameter

Wie bereits erwähnt, beträgt der konfigurierbare Maximalwert für die Warteschlangenbeschränkung in Byte 2097152. Dies entspricht etwa 18 % des gemeinsam genutzten Puffers auf ASR 920-Plattformen (11,75 MB - 45898 x 256 Byte). Wenn Sie die Warteschlangenbeschränkung jedoch in Prozent konfigurieren, können Sie bis zu 100 % erreichen. Um Prozentwerte und Bytes mit äquivalenten Werten zu vergleichen, nimmt der Test Bytewerte von 117498 bis 2097152 Bytes und Warteschlangengrenzwerte von 1 % bis 18 %:

```
queue-limit percent 1 <=> queue-limit 117498 bytes  
queue-limit percent 2 <=> queue-limit 234996 bytes  
queue-limit percent 3 <=> queue-limit 352494 bytes  
queue-limit percent 4 <=> queue-limit 469992 bytes  
queue-limit percent 5 <=> queue-limit 587490 bytes  
queue-limit percent 6 <=> queue-limit 704988 bytes  
queue-limit percent 7 <=> queue-limit 822486 bytes  
queue-limit percent 8 <=> queue-limit 939984 bytes  
queue-limit percent 9 <=> queue-limit 1057482 bytes  
queue-limit percent 10 <=> queue-limit 1174980 bytes  
queue-limit percent 11 <=> queue-limit 1292478 bytes  
queue-limit percent 12 <=> queue-limit 1409976 bytes  
queue-limit percent 13 <=> queue-limit 1527474 bytes  
queue-limit percent 14 <=> queue-limit 1644972 bytes  
queue-limit percent 15 <=> queue-limit 1762470 bytes  
queue-limit percent 16 <=> queue-limit 1879968 bytes  
queue-limit percent 17 <=> queue-limit 1997466 bytes  
queue-limit percent 18 <=> queue-limit 2097152 bytes
```

Es sind 36 Richtlinienzuweisungen konfiguriert: 18 mit Warteschlangengrenzwerten von 1 % bis 18 % und die anderen 18 mit Warteschlangengrenzwerten von 117498 bis 2097152 Byte.

```
policy-map QUEUE-LIMIT-PERCENT-X  
class DUMMY  
class class-default  
  queue-limit percent X
```

```
policy-map QUEUE-LIMIT-BYTES-X  
class DUMMY  
class class-default  
  queue-limit Y bytes
```

=> X values range from 1 to 18
=> Y values range from 117498 to 2097152

Jede Richtlinie wird mit demselben Microburst-Datenverkehr getestet, der mit IXIA generiert wird. Dieser Datenverkehr kommt an einem 10 GE-Port des ASR 920 an und wird an einem 1 GE-Port desselben Routers beendet.

Die Bursts bestehen aus 20000 Paketen mit 4 Gbit/s in Intervallen von 5 Sekunden. Die Burst-Dauer bei einer bestimmten Paketgröße:

```
1280000 bytes at 64 packet size, Burst duration: 0.00256 second
4000000 bytes at 200 packet size, Burst duration: 0.008 second
6000000 bytes at 300 packet size, Burst duration: 0.012 second
10360000 bytes at 518 packet size, Burst duration: 0.02072 second
16000000 bytes at 800 packet size, Burst duration: 0.032 second
20480000 bytes at 1024 packet size, Burst duration: 0.04096 second
28000000 bytes at 1400 packet size, Burst duration: 0.056 second
```

Testverfahren

Schritt 1: Wenden Sie die Richtlinienzuweisung QUEUE-LIMIT-BYTES-X (X=1) mit Bytes Y=117498 unter der Ausgangsschnittstelle an.

Schritt 2: Microburst-Datenverkehr 1 Minute lang ausführen.

Schritt 3: Messen Sie die Gesamtzahl der ausgegebenen Pakete und die Anzahl der verworfenen Pakete.

Schritt 4: Berechnen Sie das Verhältnis zwischen verworfenen Paketen und den insgesamt ausgegebenen Paketen.

Schritt 5: Wiederholen Sie Schritt 1. Verwenden Sie diesmal "policy-map QUEUE-LIMIT-PERCENT-X" mit dem Warteschlangenlimit-Prozentsatz X, wobei X=1 ist.

Schritt 6: Wiederholen Sie Schritt 1 mit X=X+1 für Name und Prozentwert der Richtlinienzuordnung, Y=Y+117498 Bytes. Wiederholen Sie, bis X=18 und Y=2097152 ist.

Beispiel

Messung mit der Richtlinienzuordnung QUEUE-LIMIT-BYTES-1:

```
ASR-920-1#show int Gi0/0/1
GigabitEthernet0/0/1 is up, line protocol is up
  Hardware is 24xGE-4x10GE-FIXED-S, address is 70df.2f2f.ed01 (bia 70df.2f2f.ed01)
  Internet address is 10.12.10.47/31
  MTU 8900 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
    reliability 255/255, txload 4/255, rxload 4/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1000Mbps, link type is auto, media type is T
  output flow-control is unsupported, input flow-control is on
  Carrier delay is 0 msec
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:01, output 00:00:02, output hang never
  Last clearing of "show interface" counters never
  Input queue: 0/375/0/0 (size/max/drops/flushes); Total output drops: 684263427
  Queueing strategy: Class-based queueing
```

```
Output queue: 0/40 (size/max)
30 second input rate 19475000 bits/sec, 19533 packets/sec
30 second output rate 19157000 bits/sec, 13356 packets/sec
5064106237 packets input, 4333296255278 bytes, 0 no buffer
Received 29 broadcasts (0 IP multicasts)
0 runs, 0 giants, 0 throttles
0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
0 watchdog, 726180 multicast, 0 pause input
7829367523 packets output, 4217074973677 bytes, 0 underruns
0 output errors, 0 collisions, 3 interface resets
0 unknown protocol drops
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 pause output
0 output buffer failures, 0 output buffers swapped out
```

```
ASR-920-1#show policy-map int Gi0/0/1 output
GigabitEthernet0/0/1
```

```
Service-policy output: QUEUE-LIMIT-BYTES-1
```

```
Class-map: DUMMY (match-all)
  0 packets, 0 bytes
  30 second offered rate 0000 bps
  Match: qos-group 99
```

```
Class-map: class-default (match-any)
  1044078 packets, 73085460 bytes
  30 second offered rate 9759000 bps, drop rate 0000 bps
  Match: any
```

```
queue limit 117498 bytes
(queue depth/total drops/no-buffer drops) 0/0/0
(pkts output/bytes output) 0/0
```

Messung nach 1 Minute Microburst-Datenverkehr:

```
ASR-920-1#show int Gi0/0/1
GigabitEthernet0/0/1 is up, line protocol is up
Hardware is 24xGE-4x10GE-FIXED-S, address is 70df.2f2f.ed01 (bia 70df.2f2f.ed01)
Internet address is 10.12.10.47/31
MTU 8900 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
  reliability 255/255, txload 2/255, rxload 3/255
Encapsulation ARPA, loopback not set
Keepalive set (10 sec)
Full Duplex, 1000Mbps, link type is auto, media type is T
output flow-control is unsupported, input flow-control is on
Carrier delay is 0 msec
ARP type: ARPA, ARP Timeout 04:00:00
Last input 00:00:01, output 00:00:01, output hang never
Last clearing of "show interface" counters never
Input queue: 0/375/0/0 (size/max/drops/flushes); Total output drops: 684561562
Queueing strategy: Class-based queueing
Output queue: 0/40 (size/max)
30 second input rate 13981000 bits/sec, 19643 packets/sec
30 second output rate 11256000 bits/sec, 12784 packets/sec
5064715137 packets input, 4333338878716 bytes, 0 no buffer
Received 29 broadcasts (0 IP multicasts)
0 runs, 0 giants, 0 throttles
0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
```

```

0 watchdog, 726190 multicast, 0 pause input
7829753878 packets output, 4217102018968 bytes, 0 underruns
0 output errors, 0 collisions, 3 interface resets
0 unknown protocol drops
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 pause output
0 output buffer failures, 0 output buffers swapped out

```

```

ASR-920-1#show policy-map int Gi0/0/1 output
GigabitEthernet0/0/1

```

Service-policy output: QUEUE-LIMIT-BYTES-1

```

Class-map: DUMMY (match-all)
  0 packets, 0 bytes
  30 second offered rate 0000 bps
Match: qos-group 99

```

```

Class-map: class-default (match-any)
 1847215 packets, 129305050 bytes
 30 second offered rate 10804000 bps, drop rate 0000 bps
Match: any

```

```

queue limit 117498 bytes
(queue depth/total drops/no-buffer drops) 0/387570/0
(pkts output/bytes output) 656508/45955560

```

Paketverlust-Delta: 684561562-684263427 = 298135

Gesamtpaketausgabedelta: 7829753878-7829367523 = 386355

Verhältnis zwischen Paketverlust und Paketverlust: 298135/386355 = 77 %

Testergebnisse

Wie bereits erläutert, wurden 36 Richtlinienzuweisungen getestet: 18 mit Warteschlangengrenzwerten, die zwischen 1 % und 18 % liegen, wurden mit den anderen 18 Richtlinien verglichen, die mit Warteschlangengrenzwerten konfiguriert wurden, die zwischen 117498 und 2097152 Byte liegen. Jede Richtlinienzuweisung wird mit demselben Microburst-Datenverkehr getestet, der mit IXIA generiert wurde.

In diesem Abschnitt werden die Ergebnisse dieses Tests aufgedeckt, der siebenmal wiederholt wird, um die Ergebnisse mit verschiedenen Paketgrößen in Byte zu überprüfen: 64, 200, 300, 518, 800, 1024, 1400.

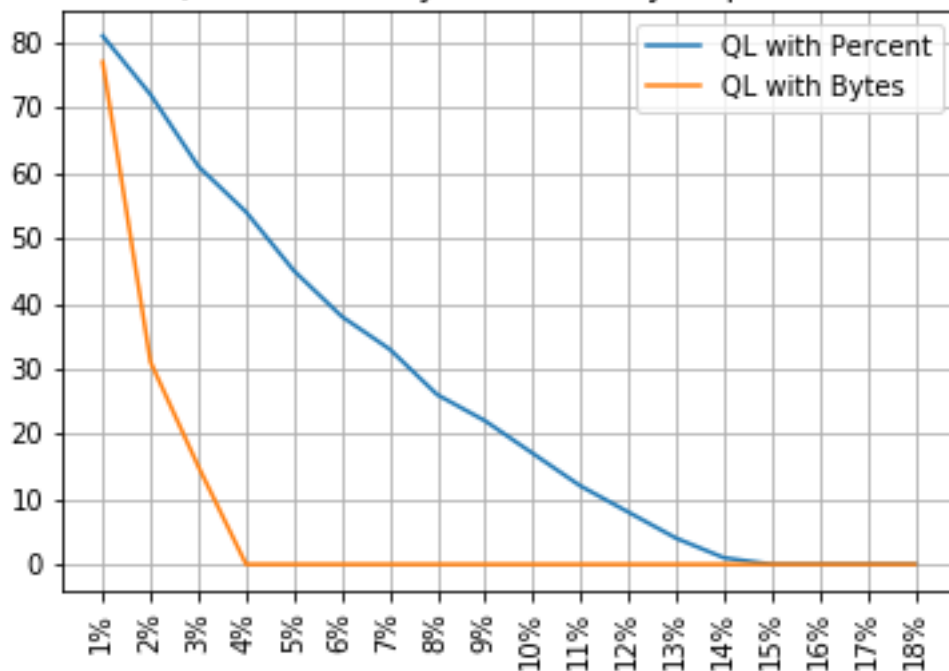
Um das Lesen zu erleichtern, werden die Ergebnisse für jede Paketgröße in einer Tabelle angezeigt und grafisch dargestellt.

QL-Prozent vs. Bytes mit 64-Byte-Paketen

Policy-map		Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent	1	81%	77%
QL Bytes/Percent	2	72%	31%

QL Bytes/Percent	3	61%	15%
QL Bytes/Percent	4	54%	0%
QL Bytes/Percent	5	45%	0%
QL Bytes/Percent	6	38%	0%
QL Bytes/Percent	7	33%	0%
QL Bytes/Percent	8	26%	0%
QL Bytes/Percent	9	22%	0%
QL Bytes/Percent	10	17%	0%
QL Bytes/Percent	11	12%	0%
QL Bytes/Percent	12	8%	0%
QL Bytes/Percent	13	4%	0%
QL Bytes/Percent	14	1%	0%
QL Bytes/Percent	15	0%	0%
QL Bytes/Percent	16	0%	0%
QL Bytes/Percent	17	0%	0%
QL Bytes/Percent	18	0%	0%

QL Percent vs bytes with 64 bytes packets

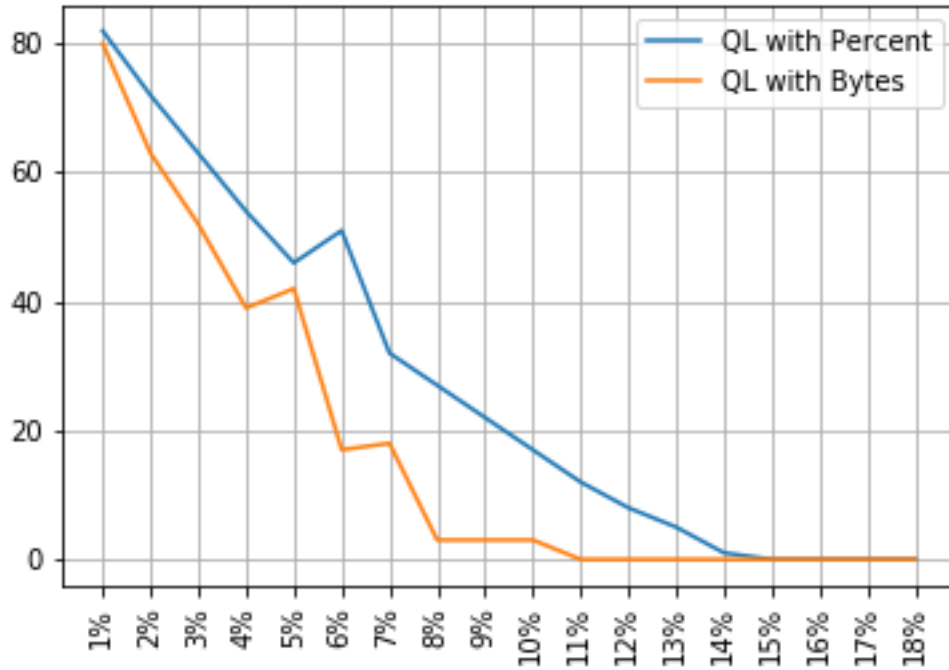


QL Prozent vs. Bytes mit 200 Bytes Packets

Policy-map	Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent 1	82%	80%
QL Bytes/Percent 2	72%	63%
QL Bytes/Percent 3	63%	52%
QL Bytes/Percent 4	54%	39%
QL Bytes/Percent 5	46%	42%
QL Bytes/Percent 6	51%	17%
QL Bytes/Percent 7	32%	18%
QL Bytes/Percent 8	27%	3%
QL Bytes/Percent 9	22%	3%
QL Bytes/Percent 10	17%	3%
QL Bytes/Percent 11	12%	0%
QL Bytes/Percent 12	8%	0%
QL Bytes/Percent 13	5%	0%
QL Bytes/Percent 14	1%	0%

QL Bytes/Percent 15	0%	0%
QL Bytes/Percent 16	0%	0%
QL Bytes/Percent 17	0%	0%
QL Bytes/Percent 18	0%	0%

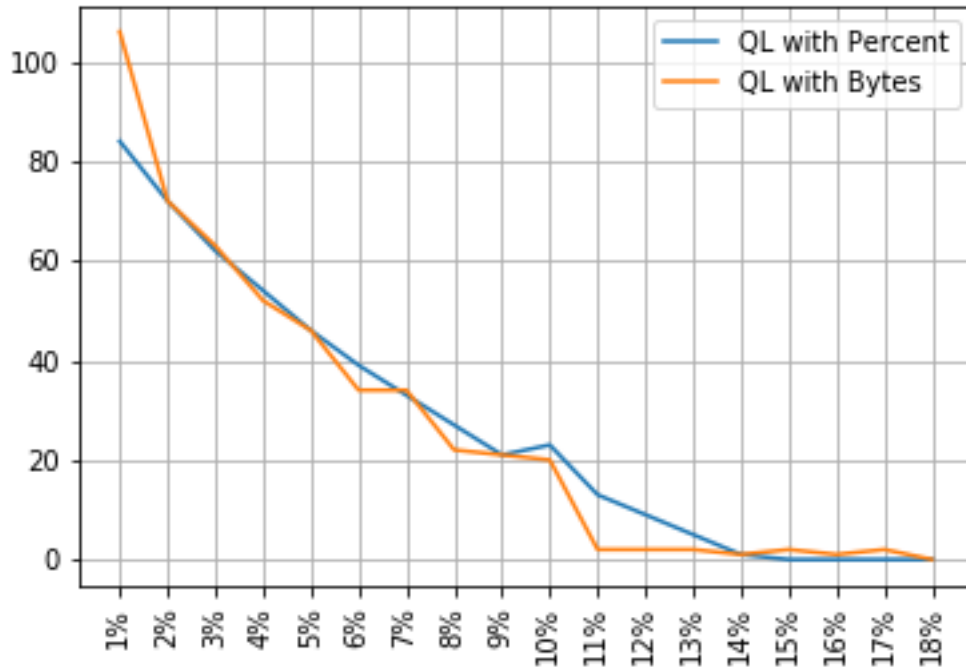
QL Percent vs bytes with 200 bytes packets



QL Prozent vs. Bytes mit 300 Bytes Packets

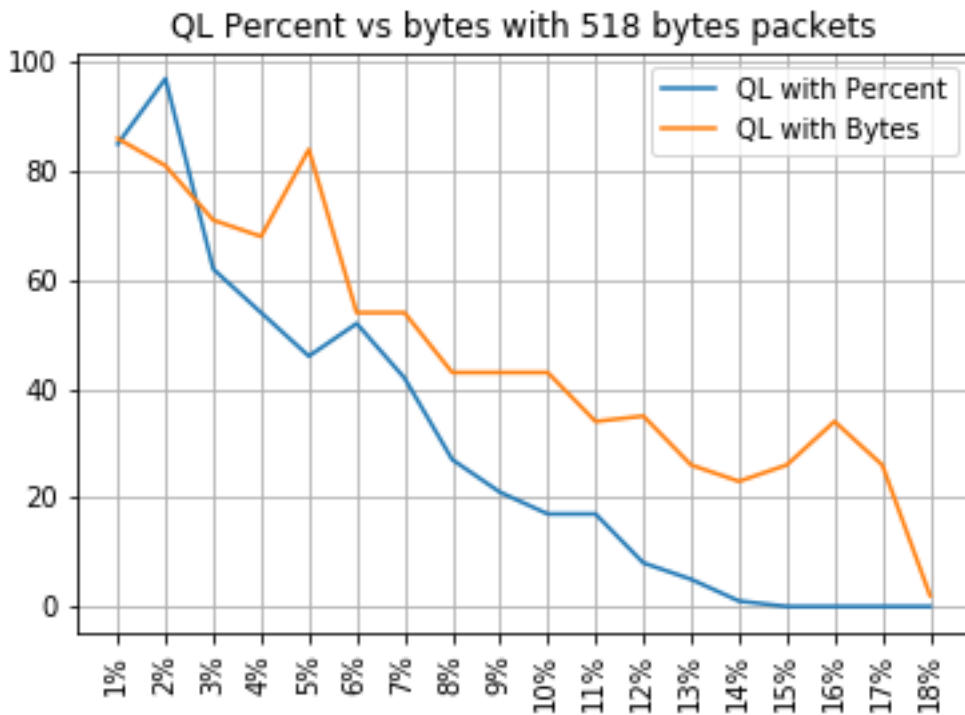
Policy-map	Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent 1	84%	106%
QL Bytes/Percent 2	72%	72%
QL Bytes/Percent 3	62%	63%
QL Bytes/Percent 4	54%	52%
QL Bytes/Percent 5	46%	46%
QL Bytes/Percent 6	39%	34%
QL Bytes/Percent 7	33%	34%
QL Bytes/Percent 8	27%	22%
QL Bytes/Percent 9	21%	21%
QL Bytes/Percent 10	23%	20%
QL Bytes/Percent 11	13%	2%
QL Bytes/Percent 12	9%	2%
QL Bytes/Percent 13	5%	2%
QL Bytes/Percent 14	1%	1%
QL Bytes/Percent 15	0%	2%
QL Bytes/Percent 16	0%	1%
QL Bytes/Percent 17	0%	2%
QL Bytes/Percent 18	0%	0%

QL Percent vs bytes with 300 bytes packets



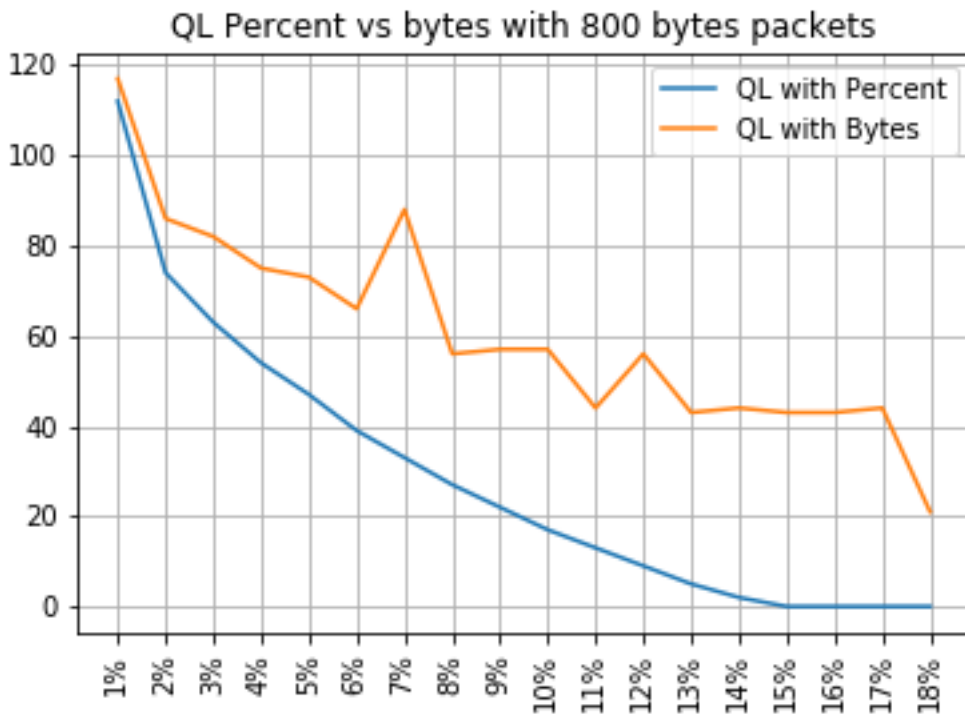
QL Prozent vs. Bytes mit 518 Bytes Packets

Policy-map	Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent 1	85%	86%
QL Bytes/Percent 2	97%	81%
QL Bytes/Percent 3	62%	71%
QL Bytes/Percent 4	54%	68%
QL Bytes/Percent 5	46%	84%
QL Bytes/Percent 6	52%	54%
QL Bytes/Percent 7	42%	54%
QL Bytes/Percent 8	27%	43%
QL Bytes/Percent 9	21%	43%
QL Bytes/Percent 10	17%	43%
QL Bytes/Percent 11	17%	34%
QL Bytes/Percent 12	8%	35%
QL Bytes/Percent 13	5%	26%
QL Bytes/Percent 14	1%	23%
QL Bytes/Percent 15	0%	26%
QL Bytes/Percent 16	0%	34%
QL Bytes/Percent 17	0%	26%
QL Bytes/Percent 18	0%	2%



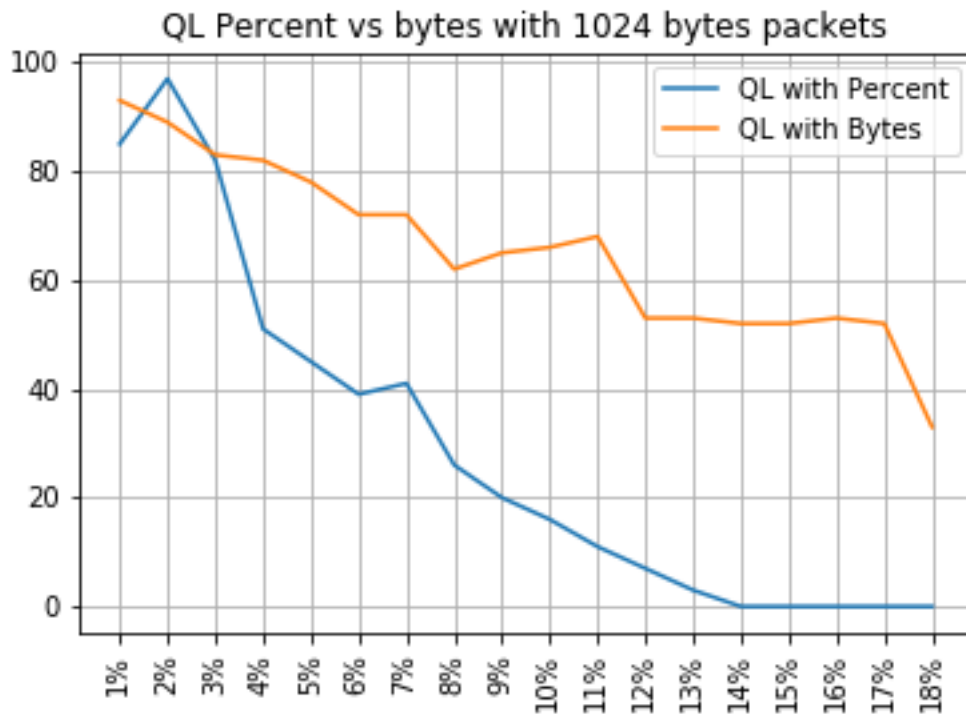
QL Prozent vs. Bytes mit 800 Bytes Packets

Policy-map	Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent 1	112%	117%
QL Bytes/Percent 2	74%	86%
QL Bytes/Percent 3	63%	82%
QL Bytes/Percent 4	54%	75%
QL Bytes/Percent 5	47%	73%
QL Bytes/Percent 6	39%	66%
QL Bytes/Percent 7	33%	88%
QL Bytes/Percent 8	27%	56%
QL Bytes/Percent 9	22%	57%
QL Bytes/Percent 10	17%	57%
QL Bytes/Percent 11	13%	44%
QL Bytes/Percent 12	9%	56%
QL Bytes/Percent 13	5%	43%
QL Bytes/Percent 14	2%	44%
QL Bytes/Percent 15	0%	43%
QL Bytes/Percent 16	0%	43%
QL Bytes/Percent 17	0%	44%
QL Bytes/Percent 18	0%	21%



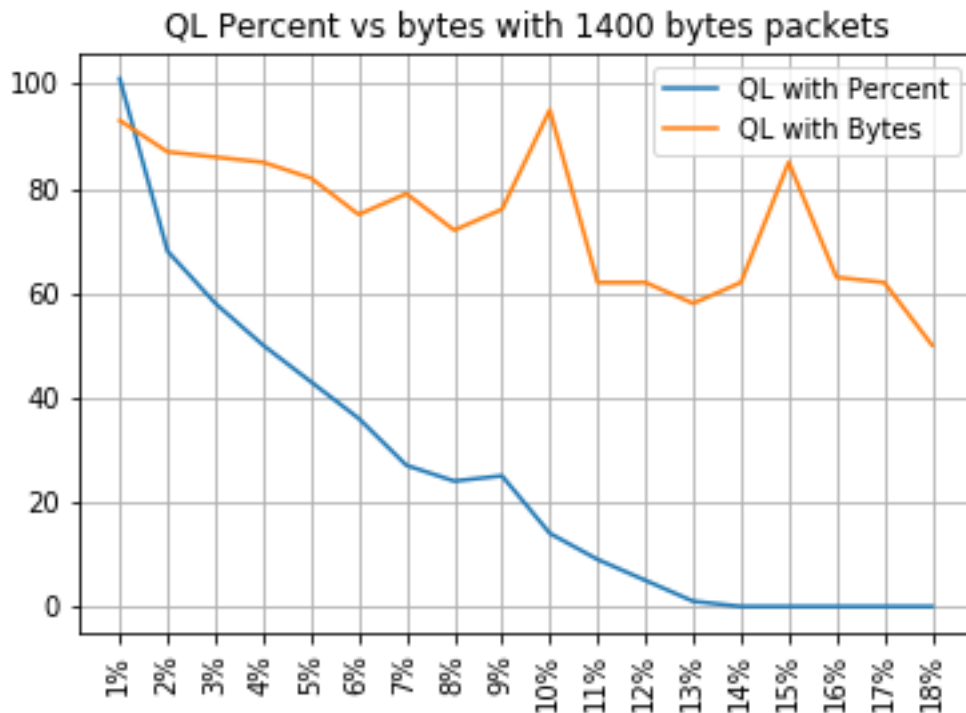
QL Prozent vs. Bytes mit 1024 Bytes Packets

Policy-map	Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent 1	85%	93%
QL Bytes/Percent 2	97%	89%
QL Bytes/Percent 3	82%	83%
QL Bytes/Percent 4	51%	82%
QL Bytes/Percent 5	45%	78%
QL Bytes/Percent 6	39%	72%
QL Bytes/Percent 7	41%	72%
QL Bytes/Percent 8	26%	62%
QL Bytes/Percent 9	20%	65%
QL Bytes/Percent 10	16%	66%
QL Bytes/Percent 11	11%	68%
QL Bytes/Percent 12	7%	53%
QL Bytes/Percent 13	3%	53%
QL Bytes/Percent 14	0%	52%
QL Bytes/Percent 15	0%	52%
QL Bytes/Percent 16	0%	53%
QL Bytes/Percent 17	0%	52%
QL Bytes/Percent 18	0%	33%



QL Prozent vs. Bytes mit 1400 Bytes Packets

Policy-map	Drop rate with PC	Drop rate with Bytes
QL Bytes/Percent 1	101%	93%
QL Bytes/Percent 2	68%	87%
QL Bytes/Percent 3	58%	86%
QL Bytes/Percent 4	50%	85%
QL Bytes/Percent 5	43%	82%
QL Bytes/Percent 6	36%	75%
QL Bytes/Percent 7	27%	79%
QL Bytes/Percent 8	24%	72%
QL Bytes/Percent 9	25%	76%
QL Bytes/Percent 10	14%	95%
QL Bytes/Percent 11	9%	62%
QL Bytes/Percent 12	5%	62%
QL Bytes/Percent 13	1%	58%
QL Bytes/Percent 14	0%	62%
QL Bytes/Percent 15	0%	85%
QL Bytes/Percent 16	0%	63%
QL Bytes/Percent 17	0%	62%
QL Bytes/Percent 18	0%	50%



Beziehung zwischen Tropfen und Paketgröße

Wie bereits erwähnt, verfügt der ASR 920 über einen internen Paketpuffer von 11,75 MB, der in 45898 Qnode mit jeweils 256 Byte aufgeteilt ist.

- Für ein Paket mit einer Größe < 256 Byte wird genau 1 Qnode verwendet
- Für ein Paket der Größe 1024 Bytes werden 4 Qnode verwendet
- Bei einem Paket mit der Größe 257 Byte werden 2 Qnode verwendet und die nicht verwendeten 255 Byte gehen verloren

Daher können Sie eine geringere Menge an großen Paketen als die Menge an kleinen Paketen speichern. Die Beziehung zwischen der Verlustrate und der Paketgröße bei entsprechenden Warteschlangengrenzwerten wird erwartet.

Differenz zwischen äquivalenten Bytes und Prozentwerten

Wie bereits erläutert, befinden sich im gemeinsamen 11,75 MB großen Puffer 45898 Qnode, auf 45900 gerundet, um die Berechnung zu vereinfachen.

Bei der Berechnung des Warteschlangengrenzwerts in Prozent wird nicht der Prozentsatz der 11,75 MB, sondern der der 45900 Qnode berechnet. Warteschlangenlimit Prozent 10 bedeutet 10% von 45900, was 4590 Qnodes ergibt.

Darüber hinaus wird der Prozentsatz der zugewiesenen Qnodes als die Anzahl der Pakete betrachtet, die unabhängig von ihrer Größe in der Warteschlange gespeichert werden können. Zurück zum vorherigen Beispiel:

queue-limit percent 10 = 4590 Qnode = 4590 Pakete.

Da diese Berechnung unabhängig von der Paketgröße ist, wird für Pakete mit einer Größe von 256 Byte oder weniger nur ein Qnode verwendet und die Gleichwertigkeit zwischen Qnode und Paketen wird beibehalten:

Warteschlangenlimit Prozent 10 = 4590 Qnode = 4590 Pakete mit 256 Bytes = 4590*256 Bytes = 1,175 MB = 10% des Puffers

Bei größeren Paketen wird jedoch ein großzügigerer Teil des Puffers zugewiesen. Dies ist z. B. die Berechnung für 1024-Byte-Pakete, bei denen jedes Paket 4 Qnode beansprucht:

Warteschlangenlimit Prozent 10 = 4.590 Qnode = 4.590 Pakete mit 1.024 Byte = 4.590*4.256 Byte = 4,7 MB = 40 % des Puffers

Vorsicht: Es wird nicht empfohlen, hohe Werte für den Warteschlangengrenzwert in Prozent zu konfigurieren.

Wenn Sie hohe Werte für den Prozentsatz der Warteschlangenbegrenzung konfigurieren, kann eine einzelne Schnittstelle vorübergehend den gesamten freigegebenen Puffer von 11,75 MB belegen.

Schlussfolgerungen

- Es ist deutlich erkennbar, dass die Effizienz von Warteschlangenlimit-Bytes bei kleinen Paketen besser ist - Warteschlangenlimit-Bytes <x> funktionieren bis zu 300 Bytes besser als Warteschlangenlimit <x>
- Bei 300-Byte-Paketen sind Warteschlangenlimit-Bytes und Warteschlangenlimit-Effizienz in Prozent identisch
- Über 300 Byte Paketgröße, Warteschlangenlimit Prozent ist effizienter. Da der Internet-Datenverkehr durchschnittlich 518 Byte beträgt, profitieren reale Szenarien mehr von der prozentualen Warteschlangenbeschränkung, wie die Kunden berichten
- Die Effizienz der Warteschlangenbegrenzung (%) verbessert sich linear in Bezug auf die Paketgröße (je größer die Pakete, desto effizienter ist die Warteschlangenbegrenzung (%) im Vergleich zu den Warteschlangenbegrenzungsbites).
- Der Warteschlangengrenzwert (%) wird als großzügiger bei der Pufferspeicherzuweisung für Pakete mit einer Größe von über 256 MB implementiert.

Wie man die Prozentwerte in realen Verkehrsszenarien annimmt

Wenn Sie Pakete mit einer Größe von 256 Byte und einem Warteschlangenlimit von 10 % haben, wissen Sie bereits, dass diese Entsprechung gültig ist:

`minimum queue-limit = 4590 Qnodes = 4590*256 bytes = 1.175 MB = 10% of the buffer`

Bei 512-Byte-Paketen ist die Auslastung nur doppelt so groß, bei 1024-Byte-Paketen ist sie viermal so groß, und so weiter.

Das bedeutet, dass die tatsächliche Warteschlangengrenze mindestens 10 % des Puffers beträgt. Wenn Sie eine maximale MTU von 1.500 Byte annehmen, benötigen Sie 6 Qnodes, um ein einzelnes Paket zu speichern. Dies ergibt eine maximale Warteschlangengrenze von:

`maximum queue-limit = 4590*6 Qnodes = 4590*256*6 bytes = 7.05 MB = 60% of the buffer`

Auf diese Weise können Sie die unteren und oberen Grenzen der Puffernutzung mit dem Warteschlangengrenzwert in Prozent 10 definieren, sodass die durchschnittliche maximale

Puffernutzung im Allgemeinen in etwa wie folgt aussieht:

```
ceil(avg_pkt_size/256)*((qlimit_percent/45900)*100)
```

Beispiel aus einem Laborgerät:

```
GigabitEthernet0/0/1 is up, line protocol is up
Hardware is 24xGE-4x10GE-FIXED-S, address is 70df.2f2f.ed01 (bia 70df.2f2f.ed01)
Internet address is 10.12.10.47/31
MTU 8900 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
reliability 255/255, txload 25/255, rxload 30/255
Encapsulation ARPA, loopback not set
Keepalive set (10 sec)
Full Duplex, 1000Mbps, link type is auto, media type is T
output flow-control is unsupported, input flow-control is on
Carrier delay is 0 msec
ARP type: ARPA, ARP Timeout 04:00:00
Last input 00:00:00, output 00:00:01, output hang never
Last clearing of "show interface" counters 00:11:43
Input queue: 0/375/0/0 (size/max/drops/flushes); Total output drops: 2036062
Queueing strategy: Class-based queueing
Output queue: 0/40 (size/max)
30 second input rate 118520000 bits/sec, 18902 packets/sec
30 second output rate 101646000 bits/sec, 16124 packets/sec
13185272 packets input, 10328798549 bytes, 0 no buffer
Received 0 broadcasts (0 IP multicasts)
0 runts, 0 giants, 0 throttles
0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
0 watchdog, 235 multicast, 0 pause input
11247114 packets output, 8870166880 bytes, 0 underruns <<< avg_pkt_size = 8870166880/11247114 =
788.66 bytes
0 output errors, 0 collisions, 0 interface resets
0 unknown protocol drops
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 pause output
0 output buffer failures, 0 output buffers swapped out
```

Sie können `avg_pkt_size` als `8870166880/11247114 ~ 788 Bytes` berechnen.

Bei einer Warteschlangenbeschränkung von 10 % beträgt die durchschnittliche maximale Puffernutzung:

```
ceil(avg_pkt_size/256)*((45900/100)*qlimit_percent)
```

Calculation example with Python:

```
>>>import math
>>>math.ceil(788/256)*((45900/100)*10)
18360.0
```

=> 18360 Qnodes = 18360 * 256 bytes = 4.7 MB = 40% of the buffer

Überprüfen der Verwendung des Microburst-Puffers

In Cisco IOS-XE-Versionen vor 16.9.3 wurde der gemeinsame Puffer der Schnittstellen für Daten- und Steuerungspakete (z. B. BFD, Routing-Protokolle, ARP, LDP, Punt-Keepalives) verwendet. Um die Verwendung des Sofortpuffers zu überprüfen, können Sie folgenden Befehl verwenden:


```
ASR-920-1#request platform software sdcli "nile bm reg buffertablefreelistcount show 0 0 0"
```

Nach 16.9.3 wurden einige Änderungen eingeführt, um die Puffernutzung zu verbessern, und es wurde in 2 aufgeteilt: 1024 Einträge (256 KB) wurden für Kontrolldatenverkehr und der Rest für Datenverkehr reserviert.

Die Puffernutzung kann in diesem Fall mit den folgenden Befehlen überwacht werden:

```
ASR-920-1#request platform software sdcli "nile bm reg
supervisorresourcereservedcounttableaccess sh 0 0 0"
reservedUsedCount = 48 (0x30)
reservedFreeCount = 976 (0x3d0)
```

```
ASR-920-1#request platform software sdcli "nile bm reg
supervisorresourcereservedcounttableaccess sh 0 2 0"
reservedUsedCount = 8114 (0x1fb2)
reservedFreeCount = 37784 (0x9398)
```

Da der Puffer Microbursts behandelt, müssen Sie den Befehl mehrmals wiederholen, um den reservierten UsedCount-Wert zu sehen, der sich von 0 unterscheidet.

Die Pufferverwendung kann einfach mit reservedUsedCount/reservedFreeCount berechnet werden, z. B. $8114/37784 = 21,5\%$ verwendet. Sobald der Burst beendet ist, muss der Puffer schnell auf 0 oder nahe daran zurückfallen.

Ab Cisco IOS-XE Version 17.6.1 ist es möglich, den gesamten Puffer sowohl für Daten- als auch für Kontrolldatenverkehr zu verwenden (vorzugsweise bei einer hohen Kontrolldatenverkehrsrate im Netzwerk) oder den Puffer wie zuvor beschrieben in 2 aufzuteilen. Die Auswahl erfolgt mit der Konfiguration dieser Anweisung (standardmäßig deaktiviert):

```
ACDC-920-1(config)#platform qos-buffer enhance enable
ACDC-920-1(config)#no platform qos-buffer enhance enable
```

Ab Cisco IOS-XE Version 17.7.1 kann die Größe, die dem Steuerungsdatenverkehr zugewiesen werden soll, auch ausgewählt werden:

```
ACDC-920-1(config)#platform qos-buffer enhance [1-4]
```

Hierbei gilt:

- 1 gibt Steuerungspuffer von 256 KB an
- 2 steht für 500 KB
- 3 steht für 756 KB
- 4 = 1 MB

Informationen zu dieser Übersetzung

Cisco hat dieses Dokument maschinell übersetzen und von einem menschlichen Übersetzer editieren und korrigieren lassen, um unseren Benutzern auf der ganzen Welt Support-Inhalte in ihrer eigenen Sprache zu bieten. Bitte beachten Sie, dass selbst die beste maschinelle Übersetzung nicht so genau ist wie eine von einem professionellen Übersetzer angefertigte. Cisco Systems, Inc. übernimmt keine Haftung für die Richtigkeit dieser Übersetzungen und empfiehlt, immer das englische Originaldokument (siehe bereitgestellter Link) heranzuziehen.