

QoS-Ausgabeplanung für Catalyst Switches der Serien 6500/6000 mit Cisco IOS-Systemsoftware

Inhalt

[Einführung](#)

[Voraussetzungen](#)

[Anforderungen](#)

[Verwendete Komponenten](#)

[Konventionen](#)

[Hintergrundinformationen](#)

[Ausgabewarteschlangen fällt aus](#)

[Ausgangswarteschlangen-Funktion verschiedener Line Cards auf dem Catalyst 6500/6000](#)

[Analyse der Warteschlangen-Funktion eines Ports](#)

[Konfiguration, Überwachung und Beispiel für die Ausgabeplanung auf dem Catalyst 6500/6000](#)

[Konfiguration](#)

[Überwachung der Ausgabeplanung und Überprüfung von Konfigurationen](#)

[Beispiel für die Ausgabeplanung](#)

[Verwenden der Ausgabeplanung zur Verringerung von Verzögerungen und Jitter](#)

[Verzögerung reduzieren](#)

[Jitter reduzieren](#)

[Schlussfolgerung](#)

[Zugehörige Informationen](#)

[Einführung](#)

Durch die Verwendung der Ausgabeplanung wird sichergestellt, dass bei starker Überbelegung kein wichtiger Datenverkehr verloren geht. In diesem Dokument werden alle Techniken und Algorithmen zur Erstellung von Ausgabevorgängen für den Catalyst 6500/6000-Switch erläutert. In diesem Dokument wird auch erläutert, wie die Ausgabeplanung auf dem Catalyst 6500/6000, auf dem die Cisco IOS® Software ausgeführt wird, konfiguriert und verifiziert wird.

Weitere Informationen zu Weighted Random Early Detection (WRED), Weighted Round Robin (WRR) und Tail Drop finden Sie unter [QoS-Ausgabeplanung für Catalyst Switches der Serien 6500 und 6000 mit CatOS-Systemsoftware](#).

[Voraussetzungen](#)

[Anforderungen](#)

Für dieses Dokument bestehen keine speziellen Anforderungen.

Verwendete Komponenten

Dieses Dokument ist nicht auf bestimmte Software- und Hardwareversionen beschränkt.

Konventionen

Weitere Informationen zu Dokumentkonventionen finden Sie unter [Cisco Technical Tips Conventions](#) (Technische Tipps zu Konventionen von Cisco).

Hintergrundinformationen

Ausgabewarteschlangen fällt aus

Ausgabeverwerfungen werden durch eine überlastete Schnittstelle verursacht. Eine häufige Ursache hierfür ist der Datenverkehr von einer Verbindung mit hoher Bandbreite, die zu einer Verbindung mit geringerer Bandbreite umgeschaltet wird, oder der Datenverkehr von mehreren eingehenden Verbindungen, die zu einer einzigen ausgehenden Verbindung geschaltet werden.

Wenn beispielsweise ein großer Teil des Datenverkehrs über eine Gigabit-Schnittstelle an eine 100-Mbit/s-Schnittstelle gesendet wird, kann dies zu einem Rückgang der Ausgabe an der 100-Mbit/s-Schnittstelle führen. Der Grund hierfür ist, dass die Ausgabewarteschlange an dieser Schnittstelle durch den übermäßigen Datenverkehr aufgrund der Geschwindigkeitsungleichheit zwischen der eingehenden und der ausgehenden Bandbreite überlastet wird. Die Datenverkehrsrate auf der Ausgangsschnittstelle kann nicht alle Pakete akzeptieren, die gesendet werden sollen.

Um das Problem zu beheben, ist die beste Lösung die Erhöhung der Leitungsgeschwindigkeit. Es gibt jedoch Möglichkeiten, Ausgabeverfahren zu verhindern, zu verringern oder zu steuern, wenn Sie die Leitungsgeschwindigkeit nicht erhöhen möchten. Ausgabeverwerfungen können nur dann verhindert werden, wenn Ausgabeverwerfen auf kurze Datenspitzen zurückzuführen sind. Wenn Ausgabetropfen durch einen konstanten Hochdatenfluss verursacht werden, können Sie diese Verwerfungen nicht verhindern. Sie können diese jedoch steuern.

Ausgangswarteschlangen-Funktion verschiedener Line Cards auf dem Catalyst 6500/6000

Wenn Sie sich über die Warteschlangenfähigkeit eines Ports nicht sicher sind, stellen Sie die **Show Queueing Interface {GigabitEthernet | fastethernet} mod/port**-Befehl. Hier sehen Sie die ersten Ausgabezeilen des Befehls **show queueing interface**. Der Port befindet sich auf einer Line Card der Supervisor Engine 1A:

```
cosmos#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy:  Weighted Round-Robin

QoS is disabled globally
Trust state: trust DSCP
Default COS is 0
Transmit group-buffers feature is enabled
Transmit queues [type = lp2q2t]:
  Queue Id      Scheduling  Num of thresholds
```

```

-----
 1          WRR low          2
 2          WRR high        2
 3          Priority         1

```

!--- Output suppressed.

Die Ausgabe zeigt, dass dieser Port über einen Ausgabewarteschlangen-Typ mit der Bezeichnung 1p2q2t verfügt.

Eine weitere Möglichkeit, den für einen bestimmten Port verfügbaren Warteschlangentyp anzuzeigen, ist der Befehl **show interface abilities (Schnittstellenfunktionen anzeigen)**:

```

la-orion#show interface gigabitethernet 6/2 capabilities
GigabitEthernet6/2
 Model:                WS-SUP720-BASE
 Type:                 No GBIC
 Speed:               1000
 Duplex:              full
 Trunk encap. type:   802.1Q,ISL
 Trunk mode:          on,off,desirable,nonegotiate
 Channel:             yes
 Broadcast suppression: percentage(0-100)
 Flowcontrol:         rx-(off,on,desired),tx-(off,on,desired)
 Membership:          static
 Fast Start:          yes
 QoS scheduling:   rx-(1p1q4t), tx-(1p2q2t)
 CoS rewrite:         yes
 ToS rewrite:         yes
 Inline power:        no
 SPAN:                source/destination
 UDLD                 yes
 Link Debounce:       yes
 Link Debounce Time: yes
 Ports on ASIC:       1-2

```

Analyse der Warteschlangen-Funktion eines Ports

Für die Catalyst 6500/6000-Switches sind mehrere Warteschlangen verfügbar. In dieser Tabelle wird die Notation der Port-QoS-Architektur erläutert:

Senden (Tx)/Empfangen (Rx)	Warteschlangenbezeichnung	Anzahl Warteschlangen	Prioritätswarteschlange	Anzahl der WRR-Warteschlangen	Anzahl und Schwellenwert für WRR-Warteschlangen
Tx	2q2t	2	—	2	2 konfigurierbare Tail Drop
Tx	1p2q2t	1	1	2	2 konfigurierbares WRED

Tx	1p3q1t	4	1	1	1 konfigurierbares WRED
Tx	1p2q1t	1	1	2	1 konfigurierbares WRED
Rx	1q4t	1	—	—	4 konfigurierbare Tail Drop
Rx	1p1q4t	2	1	1	4 konfigurierbare Tail Drop
Rx	1p1q0t	2	1	1	Nicht konfigurierbar
Rx	1p1q8t	2	1	1	8 konfigurierbare WRED
Tx	1p3q8t	4	1	1	8 konfigurierbare WRED- oder Tail-Drop-Funktion
Tx	1p7q8t	8	1	7	8 konfigurierbare WRED- oder Tail-Drop-Funktion
Rx	1q2t	1	—	—	1 konfigurierbarer Tail Drop = 1 nicht konfigurierbar
Rx	1q8t	1	—	—	8 konfigurierbare

					erbare Tail Drop
Rx	2q8t	2	—	2	8 konfiguri erbare Tail Drop

In der nächsten Tabelle sind einige Module und Warteschlangentypen auf der Rx- und Tx-Seite der Schnittstelle oder des Ports aufgeführt. Wenn Ihr Modul hier nicht aufgeführt ist, können Sie mithilfe des Befehls **show interface functions** die verfügbare Warteschlangenfunktion bestimmen. Der Befehl **show interface functions** (Schnittstellenfunktionen anzeigen) wird im Abschnitt [Output Queueing Capability von Different Line Cards auf Catalyst 6500/6000](#) beschrieben.

Modul	Rx- Warteschlangen	Tx- Warteschlangen
WS-X6K-S2-PFC2	1p1q4t	1p2q2t
WS-X6K-SUP1A-2GE	1p1q4t	1p2q2t
WS-X6K-SUP1-2GE	1q4t	2q2t
WS-X6501-10GEX4	1p1q8t	1p2q1t
WS-X6502-10GE	1p1q8t	1p2q1t
WS-X6516-GBIC	1p1q4t	1p2q2t
WS-X6516-GE-TX	1p1q4t	1p2q2t
WS-X6416-GBIC	1p1q4t	1p2q2t
WS-X6416-GE-MT	1p1q4t	1p2q2t
WS-X6316-GE-TX	1p1q4t	1p2q2t
WS-X6408A-GBIC	1p1q4t	1p2q2t
WS-X6408-GBIC	1q4t	2q2t
WS-X6524-100FX-MM	1p1q0t	1p3q1t
WS-X6324-100FX-SM	1q4t	2q2t
WS-X6324-100FX-MM	1q4t	2q2t
WS-X6224-100FX-MT	1q4t	2q2t
WS-X6548-RJ-21	1p1q0t	1p3q1t
WS-X6548-RJ-45	1p1q0t	1p3q1t
WS-X6348-RJ-21	1q4t	2q2t
WS-X6348-RJ21V	1q4t	2q2t
WS-X6348-RJ-45	1q4t	2q2t
WS-X6348-RJ-45V	1q4t	2q2t
WS-X6148-RJ-45V	1q4t	2q2t

WS-X6148-RJ21V	1q4t	2q2t
WS-X6248-RJ-45	1q4t	2q2t
WS-X6248A-TEL	1q4t	2q2t
WS-X6248-TEL	1q4t	2q2t
WS-X6024-10FL-MT	1q4t	2q2t

Konfiguration, Überwachung und Beispiel für die Ausgabeplanung auf dem Catalyst 6500/6000

Konfiguration

In diesem Abschnitt werden alle Schritte beschrieben, die erforderlich sind, um die Ausgabeplanung auf einem Catalyst 6500/6000 zu konfigurieren, auf dem die Cisco IOS Software ausgeführt wird. Informationen zur Standardkonfiguration des Catalyst 6500/6000 finden Sie in [Fall 1: QoS ist aktiviert, und](#) Abschnitt dieses Dokuments [wird ein Standardparameter verwendet](#).

Bei der Konfiguration des Catalyst 6500/6000 sind die folgenden fünf Schritte erforderlich:

1. [QoS aktivieren](#)
2. [Ordnen Sie jeden möglichen CoS-Wert einer Warteschlange und einem Schwellenwert zu](#) (optional).
3. [WRR-Gewicht konfigurieren](#) (optional)
4. [Konfigurieren der Puffer, die den einzelnen Warteschlangen zugewiesen sind](#) (optional)
5. [Konfigurieren Sie den Schwellenwert für jede Warteschlange](#) (optional).

Hinweis: Jeder dieser Schritte ist optional, mit Ausnahme von Schritt 1. Sie können entscheiden, den Standardwert für einen oder mehrere Parameter beizubehalten.

Schritt 1: QoS aktivieren

Aktivieren Sie zunächst QoS. Beachten Sie, dass QoS standardmäßig deaktiviert ist. Wenn QoS deaktiviert ist, wirkt sich die von Ihnen konfigurierte CoS-Zuordnung nicht auf das Ergebnis aus. Es gibt eine Warteschlange, die zuerst in, First out (FIFO) bereitgestellt wird, und alle Pakete werden dort verworfen.

```
cosmos#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
cosmos(config)#mls qos
```

QoS is enabled globally

```
Microflow policing is enabled globally
```

```
QoS global counters:
```

```
Total packets: 552638
IP shortcut packets: 0
Packets dropped by policing: 0
IP packets with TOS changed by policing: 0
IP packets with COS changed by policing: 0
Non-IP packets with CoS changed by policing: 0
```

Schritt 2: Ordnen Sie jeden möglichen CoS-Wert einer Warteschlange und einem Schwellenwert zu.

Weisen Sie für alle Warteschlangentypen die CoS einer Warteschlange und einem Schwellenwert zu. Die für einen 2q2t-Port-Typ definierte Zuordnung wird nicht auf einen 1p2q2t-Port angewendet. Außerdem wird die Zuordnung für 2q2t auf alle Ports mit einem 2q2t-Warteschlangenmechanismus angewendet. Stellen Sie diese **cos-map**-Befehle unter der Schnittstelle aus:

```
wrr-queue cos-map Q_number_(1-2) threshold_number_(1-2) cos_value_1 cos_value_2
priority-queue cos-map Q_number_(always 1) cos_value_1 cos_value_2
```

Hinweis: Jeder dieser Befehle muss in einer Zeile enthalten sein.

Sie können die WRR-Warteschlange separat konfigurieren. Wenn eine Prioritätswarteschlange vorhanden ist, können Sie diese mit dem Befehl **priority-queue** konfigurieren.

Hinweis: Die Warteschlangen werden immer nummeriert, beginnend mit der Warteschlange mit der niedrigsten Priorität, die möglich ist, bis hin zur Warteschlange mit der höchsten Priorität, die verfügbar ist. Beispiel:

- Warteschlange 1 ist die WRR-Warteschlange mit niedriger Priorität.
- Warteschlange 2 ist die WRR-Warteschlange mit hoher Priorität.
- Warteschlange 3 ist die Warteschlange mit höchster Priorität.

Wiederholen Sie diesen Vorgang für alle Warteschlangentypen, andernfalls bleibt die CoS-Standardzuweisung erhalten. Dies ist eine Beispielkonfiguration für 1p2q2t:

```
cosmos#configure terminal
cosmos(config)#interface gigabitethernet 1/1
cosmos(config-if)#priority-queue cos-map 1 5
!--- Assign a CoS of 5 to priority queue. cos-map configured on: Gi1/1 Gi1/2 cosmos(config-if)#wrr-queue cos-map 1 1 0 1
!--- Assign CoS 0 and 1 to the first threshold of low-priority WRR queue. cos-map configured on: Gi1/1 Gi1/2 cosmos(config-if)#wrr-queue cos-map 1 2 2 3
!--- Assign CoS 2 and 3 to the second threshold of low-priority WRR queue. cos-map configured on: Gi1/1 Gi1/2 cosmos(config-if)#wrr-queue cos-map 2 1 4 6
!--- Assign CoS 4 and 6 to the first threshold of high-priority WRR queue. cos-map configured on: Gi1/1 Gi1/2 cosmos(config-if)#wrr-queue cos-map 2 2 7
!--- Assign CoS 7 to the first threshold of high-priority WRR queue. cos-map configured on: Gi1/1 Gi1/2
```

Überprüfen Sie die Konfiguration:

```
cosmos#show queueing interface gigabitethernet 1/1
!--- Output suppressed. queue thresh cos-map ----- 1 1 0 1 1 2
2 3 2 1 4 6 2 2 7 3 1 5 !--- Output suppressed.
```

Schritt 3: WRR-Gewicht konfigurieren

Konfigurieren Sie das WRR-Gewicht für die beiden WRR-Warteschlangen. Geben Sie den folgenden Schnittstellenbefehl ein:

```
wrr-queue bandwidth weight_for_Q1 weight_for_Q2
```

Gewicht 1 bezieht sich auf die Warteschlange 1, die die WRR-Warteschlange mit niedriger Priorität sein sollte. Halten Sie dieses Gewicht immer einen Wert unter Gewicht 2. Das Gewicht kann zwischen 1 und 255 liegen. Verwenden Sie diese Formeln, um den Prozentsatz zuzuweisen:

- Zur Warteschlange 1— $[\text{Gewicht 1} / (\text{Gewicht 1} + \text{Gewicht 2})]$
- Zur Warteschlange 2— $[\text{Gewicht 2} / (\text{Gewicht 1} + \text{Gewicht 2})]$

Sie müssen das Gewicht für alle Warteschlangentypen definieren. Diese Gewichtstypen müssen nicht identisch sein. Dies ist ein Beispiel für 2q2t, bei dem Warteschlange 1 20 Prozent der Zeit und Warteschlange 2 80 Prozent der Zeit serviert wird:

```
cosmos#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
cosmos(config)#interface gigabitethernet 1/1
cosmos(config-if)#wrr-queue bandwidth ?
<1-255> enter bandwidth weight between 1 and 255
cosmos(config-if)#wrr-queue bandwidth 20 80
!--- Queue 1 is served 20% of the time, and queue 2 is served !--- 80% of the time.
cosmos(config-if)#
```

Überprüfen Sie die Konfiguration:

```
cosmos#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Port is untrusted
Default cos is 0
Transmit queues [type = 1p2q2t]:
  Queue Id      Scheduling  Num of thresholds
  -----
  1             WRR low    2
  2             WRR high   2
  3             Priority   1

  WRR bandwidth ratios:  20[queue 1]  80[queue 2]
  queue-limit ratios:    90[queue 1]  5[queue 2]
!--- Output suppressed.
```

Hinweis: Sie können verschiedene WRR-Gewichtungen für jede Schnittstelle konfigurieren, wenn die CatOS-Software nicht verwendet werden kann.

Schritt 4: Konfigurieren der Puffer, die den einzelnen Warteschlangen zugewiesen sind

Sie müssen das Verhältnis der Übertragungswarteschlange definieren. Dadurch wird festgelegt, wie die Puffer auf die verschiedenen Warteschlangen aufgeteilt werden.

```
wrr-queue queue-limit percentage_WRR_Q1 percentage_WRR_Q2
cosmos(config)#interface gigabitethernet 1/2
cosmos(config-if)#wrr-queue queue-limit 70 15
!--- Queue 1 has 70% of the buffers. !--- Queues 2 and 3 both have 15% of the buffers. queue-
limit configured on: Gi1/1 Gi1/2
```

Hinweis: Wenn die Warteschlangenfunktion des Gigabit-Ports 1p1q2t beträgt, müssen Sie

dieselbe Ebene für die Warteschlange mit höchster Priorität und für die WRR-Warteschlange mit hoher Priorität verwenden. Diese Stufen können aus Hardwaregründen nicht voneinander abweichen. Es wird nur die Bandbreite für die beiden WRR-Warteschlangen konfiguriert. Sie verwenden automatisch denselben Wert für die WRR-Warteschlange mit hoher Priorität und die Warteschlange mit strikter Priorität, sofern vorhanden.

Einige Warteschlangentypen verfügen nicht über eine einstellbare Warteschlangengröße. Ein Beispiel ist 1p3q1t, das auf WS-X6548RJ45 verfügbar ist. Diese Warteschlangentypen sind fest, und Sie können sie nicht ändern.

Überprüfen Sie die Konfiguration:

```
cosmos#show queueing interface gigabitethernet 1/2
Interface GigabitEthernet1/2 queueing strategy:  Weighted Round-Robin
Port QoS is enabled
Port is untrusted
Default cos is 0
Transmit queues [type = 1p2q2t]:
  Queue Id      Scheduling  Num of thresholds
  -----
     1          WRR low           2
     2          WRR high           2
     3          Priority           1

WRR bandwidth ratios:      5[queue 1] 255[queue 2]
queue-limit ratios:       70[queue 1] 15[queue 2]
```

Hinweis: Es ist am besten, den größten Teil der Puffer für die WRR-Warteschlange mit niedriger Priorität zu belassen. In dieser Warteschlange müssen Sie zusätzliche Pufferung aktivieren. Die anderen Warteschlangen werden mit höherer Priorität bedient.

[Schritt 5: Konfigurieren der Schwellenwertstufe für jede Warteschlange](#)

Konfigurieren Sie als letzten Schritt den Schwellenwert für die WRED-Warteschlange oder für die Tail-Drop-Warteschlange. Diese Liste enthält die Befehle:

- Führen Sie für Warteschlangen, die WRED als Drop-Mechanismus für den Schwellenwert verwenden, folgende Befehle aus:

```
wrr-queue random-detect min-threshold Q_number threshold_1_value threshold_2_value
wrr-queue random-detect max-threshold Q_number threshold_1_value threshold_2_value
```

Hinweis: Jeder dieser Befehle muss in einer Zeile enthalten sein.

- Geben Sie für Warteschlangen, die Tail Drop als Drop-Mechanismus verwenden, den folgenden Befehl aus:

```
wrr-queue threshold Q_number threshold_1_value threshold_2_value
```

Hinweis: Dieser Befehl sollte in einer Zeile stehen.

Konfiguration für eine WRED-Warteschlange:

```
cosmos(config)#interface gigabitethernet 1/1
cosmos(config-if)#wrr-queue random-detect min-threshold 1 20 50
```

```

!--- This sets the threshold of queue 1 to 20 and 50% minimum threshold !--- configured on Gi1/1
Gi1/2. cosmos(config-if)#wrr-queue random-detect min-threshold 2 20 50
!--- This sets the threshold of queue 2 to 20 and 50% minimum threshold !--- configured on Gi1/1
Gi1/2. cosmos(config-if)#wrr-queue random-detect max-threshold 1 50 80
!--- This sets the threshold of queue 1 to 50 and 80% maximum threshold !--- configured on Gi1/1
Gi1/2. cosmos(config-if)#wrr-queue random-detect max-threshold 2 40 60
!--- This sets the threshold of queue 2 to 49 and 60% maximum threshold !--- configured on Gi1/1
Gi1/2.

```

Konfiguration für eine Tail-Drop-Warteschlange:

```

cosmos(config)#interface fastethernet 3/1
cosmos(config-if)#wrr-queue threshold ?
<1-2> enter threshold queue id (1-2)
cosmos(config-if)#wrr-queue threshold 1 ?
<1-100> enter percent of queue size between 1 and 100
cosmos(config-if)#wrr-queue threshold 1 50 100
!--- This sets the tail drop threshold for this 2q2t interface for !--- queue 1 (low-priority)
to 50 and 100% of the buffer. threshold configured on: Fa3/1 Fa3/2 Fa3/3 Fa3/4 Fa3/5 Fa3/6 Fa3/7
Fa3/8 Fa3/9 Fa3/10 Fa3/11 Fa3/12 cosmos(config-if)# cosmos(config-if)# cosmos(config-if)#wrr-
queue threshold 2 40 100
!--- This sets the tail drop threshold for this 2q2t interface for !--- queue 2 (high-priority)
to 40 and 100% of the buffer. threshold configured on: Fa3/1 Fa3/2 Fa3/3 Fa3/4 Fa3/5 Fa3/6 Fa3/7
Fa3/8 Fa3/9 Fa3/10 Fa3/11 Fa3/12 cosmos(config-if)#

```

Überprüfen Sie die Konfiguration:

```

cosmos#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Port is untrusted
Default cos is 0
Transmit queues [type = lp2q2t]:
Queue Id      Scheduling  Num of thresholds
-----
1             WRR low    2
2             WRR high   2
3             Priority    1

WRR bandwidth ratios: 20[queue 1] 80[queue 2]
queue-limit ratios: 70[queue 1] 15[queue 2]

queue random-detect-min-thresholds
-----
1      20[1] 50[2]
2      20[1] 50[2]

queue random-detect-max-thresholds
-----
1      50[1] 80[2]
2      40[1] 60[2]

```

```

cosmos#show queueing interface fastethernet 3/1
Interface FastEthernet3/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Port is untrusted
Default cos is 0
Transmit queues [type = 2q2t]:
Queue Id      Scheduling  Num of thresholds
-----
1             WRR low    2

```

```

2          WRR high          2

WRR bandwidth ratios: 100[queue 1] 255[queue 2]
queue-limit ratios:   90[queue 1] 10[queue 2]

queue tail-drop-thresholds
-----
1      50[1] 100[2]
2      40[1] 100[2]

```

Sie können den Grenzwert nicht konfigurieren und der Warteschlange pro Port die CoS zuweisen. Alle Änderungen werden auf eine Reihe zusammenhängender Ports angewendet:

- Vier Ports für Gigabit Line Cards - die Ports 1 bis 4 sind miteinander verbunden, die Ports 5 bis 8 sind miteinander verbunden.
- Zwölf Ports für 10/100-Ports oder 100 Glasfaserports basierend auf 1q4t/2q2t-Warteschlangen - 1 bis 12, 13 bis 24, 25 bis 36 und 36 bis 48.
- Um den genauen Port zu dem gleichen ASIC zu bestimmen, verwenden Sie den Befehl **show interface abilities (Schnittstellenfunktionen anzeigen)**.

Überwachung der Ausgabeplanung und Überprüfung von Konfigurationen

Der einfachste Befehl, der zur Überprüfung der aktuellen Laufzeitkonfiguration eines Ports hinsichtlich der Ausgabeplanung ausgeführt werden kann, ist die **Show Queueing Interface {GigabitEthernet | fastethernet} slot/port**-Befehl. Dieser Befehl zeigt die Warteschlangenart am Port, die Zuordnung von CoS zu den verschiedenen Warteschlangen und Schwellenwerten, die Pufferfreigabe und das WRR-Gewicht an. Hier beträgt der WRR von 20 % für die Warteschlange 1 und der WRR von 80 % für die Warteschlange 2. Der Befehl zeigt außerdem alle konfigurierten Informationen für die Ausgabeplanung sowie die Anzahl der Pakete an, die in jeder Warteschlange für jeden Schwellenwert verworfen werden:

```

cosmos#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Port is untrusted
Default COS is 0
Transmit queues [type = lp2q2t]:
  Queue Id      Scheduling  Num of thresholds
  -----
  1             WRR low    2
  2             WRR high   2
  3             Priority   1
WRR bandwidth ratios: 20[queue 1] 80[queue 2]
queue-limit ratios:   70[queue 1] 15[queue 2]

queue random-detect-max-thresholds
-----
1      50[1] 80[2]
2      40[1] 60[2]

queue thresh cos-map
-----
1      1      0 1
1      2      2 3
2      1      4 6
2      2      7

```

3 1 5

Receive queues [type = 1plq4t]:

Queue Id	Scheduling	Num of thresholds
1	Standard	4
2	Priority	1

queue tail-drop-thresholds

1	100[1]	100[2]	100[3]	100[4]
---	--------	--------	--------	--------

queue thresh cos-map

1	1	0	1
1	2	2	3
1	3	4	
1	4	6	7
2	1	5	

Packets dropped on Transmit:

BPDU packets: 0

queue	thresh	dropped	[cos-map]
1	1	0	[0 1]
1	2	0	[2 3]
2	1	0	[4 6]
2	2	0	[7]
3	1	0	[5]

Packets dropped on Receive:

BPDU packets: 0

queue	thresh	dropped	[cos-map]
1	1	0	[0 1]
1	2	0	[2 3]
1	3	0	[4]
1	4	0	[6 7]
2	1	0	[5]

Beispiel für die Ausgabeplanung

Dieser Datenverkehr wird auf den Catalyst 6500/6000 eingespeist:

- Gigabit 1/2 im Port: ein Gigabit Datenverkehr mit Priorität Null
- Gigabit 5/2 im Port: 133 MB Datenverkehr mit einer Priorität von sieben
133 MB Datenverkehr mit einer Priorität von sechs
133 MB Datenverkehr mit einer Priorität von fünf
133 MB Datenverkehr mit einer Priorität von vier
133 MB Datenverkehr mit Vorrang von drei
133 MB Datenverkehr mit Vorrang von zwei
133 MB Datenverkehr mit einer Priorität von einem

Der gesamte Unicast-Datenverkehr verlässt den Switch pro Port Gigabit 1/1, was sehr überbelegt ist.

Fall 1: QoS ist aktiviert, und es wird ein Standardparameter verwendet.

Der Befehl `show queueing interface gigabitethernet 1/1` konfiguriert die gesamte Ausgabe in

diesem Beispiel. Der Befehl enthält zusätzliche Informationen zur Eingabeplanung. Da dieses Dokument jedoch nur die Ausgabeplanung behandelt, wird diese Ausgabe unterdrückt.

Wenn QoS global aktiviert ist und alle Standardparameter verwendet werden, ergibt sich diese Ausgabe nach einigen Minuten:

```
nelix#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Trust state: trust DSCP
Default cos is 0
Transmit queues [type = lp2q2t]:
  Queue Id      Scheduling  Num of thresholds
  -----
  1             WRR low     2
  2             WRR high    2
  3             Priority    1

WRR bandwidth ratios: 100[queue 1] 255[queue 2]
queue-limit ratios:   90[queue 1]   5[queue 2]

queue random-detect-max-thresholds
-----
  1    40[1] 100[2]
  2    40[1] 100[2]

queue thresh cos-map
-----
  1    1    0 1
  1    2    2 3
  2    1    4
  2    2    6 7
  3    1    5

Packets dropped on Transmit:
  BPDU packets: 0

queue thresh      dropped  [cos-map]
-----
  1    1    149606424  [0 1 ]
  1    2           0  [2 3 ]
  2    1    16551394  [4 ]
  2    2    4254446   [6 7 ]
  3    1           0  [5 ]
```

In dieser Ausgabe sind die Standardwerte:

- WRR-Gewicht für Warteschlange 1-100 / (100 + 255) = 28 %
- WRR-Gewicht für Warteschlange 2-255 / (255 + 100) = 72 %
- Puffer-Freigabe: - 90 % für Warteschlange 1, 5 % für Warteschlange 2 und 5 % für Warteschlangen mit strikter Priorität

Die meisten Pakete in der WRR-Warteschlange mit niedriger Priorität werden verworfen, aber einige werden trotzdem für beide Schwellenwerte in der WRR-Warteschlange mit hoher Priorität verworfen. Insgesamt gibt es 170.412.264 Tropfen (149.606.424 + 16.551.394 + 4.254.446).

Diese Tropfen sind wie folgt geteilt:

- 149.606.424 / 170.412.264 = 88 % der Verwerfungen in der Warteschlange 1 (erstes Schwellenwertpaket mit CoS 0 und 1)

- $16.551.394 / 170.412.264 = 10\%$ der Verwerfungen in der Warteschlange 2 (erstes Schwellenwertpaket mit CoS 4)
- $4.254.446 / 170.412.264 = 2\%$ der Verwerfungen in der Warteschlange 2 (zweites Schwellenwertpaket mit CoS von 6 oder 7)

Hinweis: In der Warteschlange mit strikter Priorität werden keine Verluste angezeigt.

Fall 2: WRR-Gewicht ändern

Wie in [Fall 1](#) erwähnt: [QoS ist aktiviert und ein Abschnitt mit einem Standardparameter verwendet](#), Pakete in der Warteschlange 2 werden noch verworfen. Ändern Sie das WRR-Gewicht, um mehr Bandbreite für die Warteschlange 2 bereitzustellen. Nun ist Warteschlange 1 4 Prozent der Zeit leer und Warteschlange 2 96 Prozent der Zeit leer:

```
show run interface gigabitethernet 1/1
interface GigabitEthernet1/1
no ip address
wrr-queue bandwidth 10 255
mls qos trust dscp
switchport
switchport mode access
end
```

```
nelix#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Trust state: trust DSCP
Default cos is 0
Transmit queues [type = lp2q2t]:
Queue Id      Scheduling  Num of thresholds
-----
1             WRR low    2
2             WRR high   2
3             Priority   1

WRR bandwidth ratios:  10[queue 1] 255[queue 2]
queue-limit ratios:    90[queue 1]  5[queue 2]

queue random-detect-max-thresholds
-----
1    40[1] 100[2]
2    40[1] 100[2]

queue thresh cos-map
-----
1    1    0 1
1    2    2 3
2    1    4
2    2    6 7
3    1    5

Packets dropped on Transmit:
BPDU packets: 0
```

```
queue thresh      dropped  [cos-map]
-----
1    1          2786205  [0 1 ]
1    2              0  [2 3 ]
```

```

2      1      11363 [4 ]
2      2      69 [6 7 ]
3      1      0 [5 ]

```

Wie in dieser Ausgabe zu sehen ist, ist der Prozentsatz an Verlusten in Warteschlange 2 jetzt viel niedriger. Insgesamt sind 2.797.637 Tropfen auf diese Weise aufgeteilt:

- $2.786.205 / 2.797.637 = 99,591$ % der Drops in Warteschlange 1 (mit Paket von CoS 0 und 1)
- $11.363 / 2.797.637 = 0,408$ % der Verwerfungen in der Warteschlange 2 (erster Grenzwert bei Paket-CoS 4)
- $69/2.797.637 = 0,001$ % der Verwerfen in Warteschlange 2 (zweiter Grenzwert für Paket mit CoS 6 und 7)

Wenn Sie verschiedene WRR-Gewichtungen verwenden, wird mehr QoS in Warteschlange 2 sichergestellt.

Fall 3: Zusätzliche WRR-Gewichtsänderung

Mit dem WRR-Gewicht können Sie sogar noch aggressiver vorgehen. In dieser Beispielausgabe werden nur 0,39 Prozent des Gewichts an die Warteschlange 1 abgegeben:

```
show run interface gigabitethernet 1/1
```

```

interface GigabitEthernet1/1
no ip address
wrr-queue bandwidth 1 255
mls qos trust dscp
switchport
switchport mode access
end

```

```
nelix#show queueing interface gigabitethernet 1/1
```

```
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
```

```
Port QoS is enabled
```

```
Trust state: trust DSCP
```

```
Default cos is 0
```

```
Transmit queues [type = lp2q2t]:
```

```
Queue Id      Scheduling  Num of thresholds
```

```
-----
```

```
1            WRR low      2
```

```
2            WRR high     2
```

```
3            Priority     1
```

```
WRR bandwidth ratios: 1[queue 1] 255[queue 2]
```

```
queue-limit ratios: 90[queue 1] 5[queue 2]
```

```
queue random-detect-max-thresholds
```

```
-----
```

```
1 40[1] 100[2]
```

```
2 40[1] 100[2]
```

```
queue thresh cos-map
```

```
-----
```

```
1 1 0 1
```

```
1 2 2 3
```

```
2 1 4
```

```
2 2 6 7
```

```
3 1 5
```

```
Packets dropped on Transmit:
```

BPDU packets: 0

```
queue thresh    dropped  [cos-map]
-----
1      1      2535315  [0 1 ]
1      2           0  [2 3 ]
2      1       705  [4 ]
2      2        73  [6 7 ]
3      1           0  [5 ]
```

Selbst bei einem aggressiven WRR-Gewicht werden Pakete in der Warteschlange 2 immer noch verworfen. Im Vergleich dazu sind es jedoch nicht viele Pakete. In der Warteschlange 2 treten jetzt nur noch 0,03 Prozent Pakete verloren.

Fall 4: Ändern der Warteschlangenbegrenzungs-Pufferzuweisung

Wie im [Fall 2](#) zu sehen: [WRR-Gewicht](#) und [Fall 3 ändern: Zusätzliche](#) Abschnitte zur [WRR-Gewichtsmodifizierung](#), Pakete fallen immer noch in der Warteschlange 2 ab, obwohl der WRR-Prozentsatz Ihnen versichert, dass der Wert minimal ist. Wenn jedoch der zweite Schwellenwert (auf 100 Prozent festgelegt) in der Warteschlange 2 erreicht wird, werden einige Pakete weiterhin verworfen.

Um dies zu verbessern, ändern Sie die Warteschlangengrenze (Größe des Puffers, der jeder Warteschlange zugewiesen ist). In diesem Beispiel ist der Warteschlangenlimit für Warteschlange 1 auf 70 %, für Warteschlange 2 auf 15 % und für Warteschlangen mit strikter Priorität auf 15 % festgelegt:

```
show run gigabitethernet 1/1
interface GigabitEthernet1/1
no ip address
wrr-queue bandwidth 1 255
wrr-queue queue-limit 70 15
mls qos trust dscp
switchport
switchport mode access
end
```

```
nelix#show queueing interface gigabitethernet 1/1
Interface GigabitEthernet1/1 queueing strategy: Weighted Round-Robin
Port QoS is enabled
Trust state: trust DSCP
Default cos is 0
Transmit queues [type = lp2q2t]:
Queue Id    Scheduling  Num of thresholds
-----
1           WRR low    2
2           WRR high  2
3           Priority   1

WRR bandwidth ratios:    1[queue 1] 255[queue 2]
queue-limit ratios:     70[queue 1] 15[queue 2]

queue random-detect-max-thresholds
-----
1      40[1] 100[2]
2      40[1] 100[2]

queue thresh cos-map
```



```

-----
1      1      0 1
1      2      2 3
2      1      4
2      2      6 7
3      1      5

```

```

Receive queues [type = 1plq4t]:
Queue Id      Scheduling  Num of thresholds
-----
1             Standard    4
2             Priority    1

```

```

queue tail-drop-thresholds
-----
1      100[1] 100[2] 100[3] 100[4]

```

```

queue thresh cos-map
-----
1      1      0 1
1      2      2 3
1      3      4
1      4      6 7
2      1      5

```

```

Packets dropped on Transmit:
BPDU packets: 0

```

```

queue thresh      dropped  [cos-map]
-----
1      1          154253046  [0 1 ]
1      2              0  [2 3 ]
2      1              0  [4 ]
2      2              0  [6 7 ]
3      1              0  [5 ]

```

Die Drops treten nun nur in der Warteschlange 1 auf.

[Verwenden der Ausgabeplanung zur Verringerung von Verzögerungen und Jitter](#)

Die Fallstudien im Abschnitt [Beispiel für die Ausgabeplanung](#) veranschaulichen die Vorteile der Implementierung der Ausgabeplanung, um einen Rückgang des VoIP- oder geschäftskritischen Datenverkehrs bei Überbelegung des Ausgangsports zu vermeiden. Eine Überbelegung tritt in einem normalen Netzwerk (insbesondere bei einer Gigabit-Verbindung) nicht sehr häufig auf. Eine Überbelegung sollte nur zu Spitzenzeiten des Datenverkehrs oder bei Datenverkehrsspitzen innerhalb kürzester Zeit erfolgen.

Auch ohne Überbelegung kann die Ausgabeplanung in einem Netzwerk, in dem die QoS durchgängig implementiert wird, von großem Nutzen sein. In diesem Abschnitt finden Sie Beispiele dafür, wie die Ausgabeplanung Verzögerungen und Jitter reduzieren kann.

[Verzögerung reduzieren](#)

Die Verzögerung eines Pakets nimmt zu, da die Zeit im Puffer jedes Switches "verloren" ist, während das Paket auf die Übertragung wartet. Beispielsweise wird ein kleines Sprachpaket mit einer CoS von 5 während einer großen Datensicherung oder Dateiübertragung von einem Port gesendet. Angenommen, es gibt keine QoS für den Ausgangsport, und das kleine Sprachpaket

wird nach 10 großen 1500-Byte-Paketen in die Warteschlange gestellt. In diesem Fall können Sie leicht berechnen, dass die für die Übertragung der 10 großen Pakete benötigte Gigabit-Geschwindigkeit folgendermaßen aussieht:

- $(10 \times 1500 \times 8) = 120.000$ Bit übertragen in 120 Mikrosekunden

Wenn dieses Paket acht oder neun Switches durchlaufen muss, während es das Netzwerk durchläuft, kann es zu einer Verzögerung von etwa 1 Millisekunde kommen. Dies umfasst nur Verzögerungen in der Ausgabewarteschlange des Switches, der im Netzwerk überquert wird.

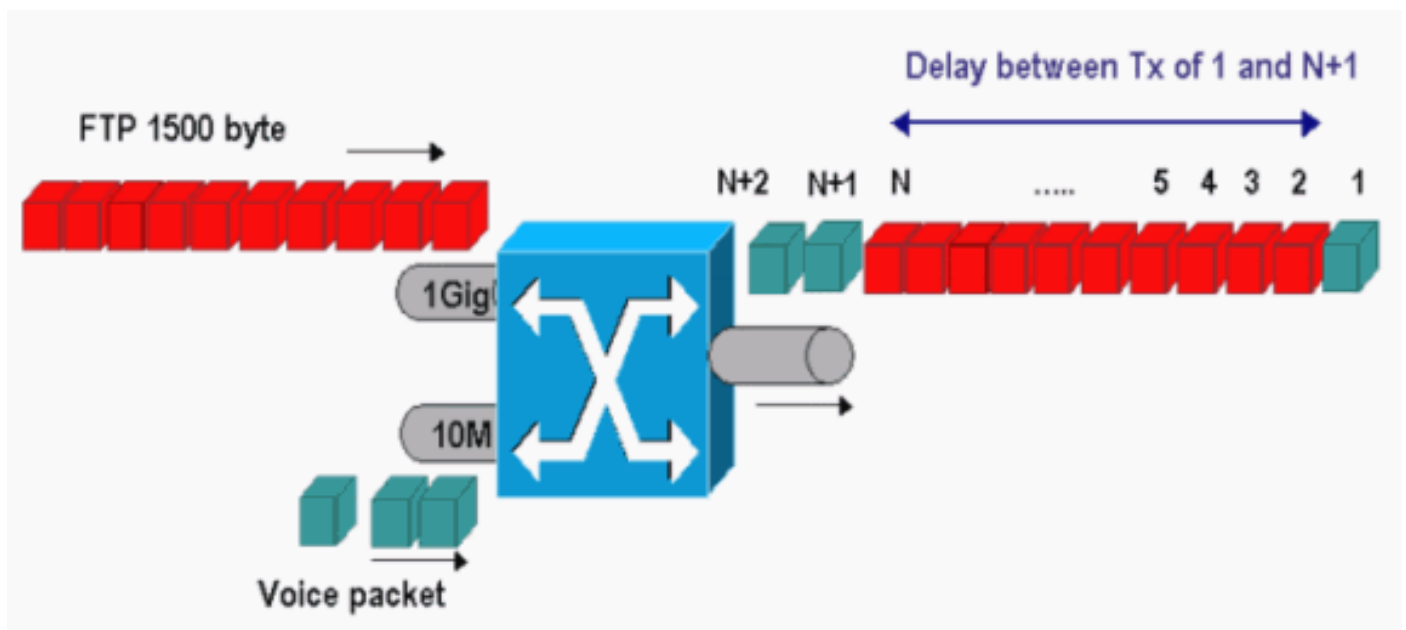
Hinweis: Wenn Sie dieselben 10 großen Pakete auf einer 10-MB-Schnittstelle in die Warteschlange stellen müssen (z. B. an ein IP-Telefon und einen PC angeschlossen), wird folgende Verzögerung angezeigt:

- $(10 \times 1500 \times 8) = 120.000$ Bit übertragen in 12 Millisekunden

Die Implementierung der Ausgabeplanung stellt sicher, dass Sprachpakete mit einer CoS von 5 in die Warteschlange mit höchster Priorität gestellt und vor allen Paketen mit einer CoS von weniger als 5 gesendet werden. Dadurch werden Verzögerungen reduziert.

Jitter reduzieren

Ein weiterer wichtiger Vorteil der Produktionsplanung ist die Reduzierung von Jitter. Bei Jitter handelt es sich um die Verzögerungsvariante für Pakete im selben Fluss. Dieses Beispielszenario zeigt, wie die Ausgabeplanung Jitter reduzieren kann:



In diesem Szenario muss derselbe Ausgangsport zwei Streams senden:

- Ein eingehender Sprach-Stream auf einem 10-MB-Ethernet-Port.
- Ein eingehender FTP-Stream auf einem 1-Gigabit-Ethernet-Uplink-Port.

Beide Streams verlassen den Switch über denselben Ausgangsport. In diesem Beispiel wird veranschaulicht, was ohne die Verwendung der Ausgabeplanung geschehen kann. Alle großen Datenpakete können zwischen zwei Sprachpaketen verschoben werden. Dadurch entsteht Jitter beim Empfang des Sprachpakets vom gleichen Stream. Es besteht eine größere Verzögerung zwischen dem Empfang von Paket 1 und Paket $n + 1$, wenn der Switch das große Datenpaket überträgt. Die Verzögerung zwischen $n + 1$ und $n + 2$ ist jedoch vernachlässigbar. Dies führt zu

Jitter im Sprachdatenverkehrsstrom. Sie können dieses Problem mit einer Warteschlange mit strikter Priorität leicht vermeiden. Stellen Sie sicher, dass Sie den CoS-Wert der Sprachpakete der Warteschlange mit strikter Priorität zuordnen.

Schlussfolgerung

In diesem Dokument finden Sie Fallstudien zur Konfiguration und Fehlerbehebung von Ausgabewarteschlangenplanungen auf einem Catalyst 6500/6000, auf dem Cisco IOS Software ausgeführt wird. Sie haben auch die Vorteile der Ausgabeplanung in den meisten Netzwerken mit Sprachdatenverkehr erkannt:

- Verhindert den Verlust von kritischem Datenverkehr bei Überbelegung des Ausgabeports.
- Reduziert die Verzögerung.
- Reduziert den Jitter.

Zugehörige Informationen

- [QoS-Ausgabeplanung für Catalyst Switches der Serien 6500 und 6000 mit CatOS-Systemsoftware](#)
- [Quality of Service auf Catalyst Switches der Serie 6000](#)
- [Support-Seiten für LAN-Produkte](#)
- [Support-Seite für LAN-Switching](#)
- [Technischer Support und Dokumentation - Cisco Systems](#)