

目次

[概要](#)

[前提条件](#)

[要件](#)

[使用するコンポーネント](#)

[表記法](#)

[Catalyst 3550 スイッチでのポートの出力キューイング機能](#)

[ギガビット ポートと非ギガビット ポートの両方がサポートする機能](#)

[ギガビット ポートのみがサポートする機能](#)

[非ギガビット ポートのみがサポートする機能](#)

[キューに対する CoS のマッピング](#)

[完全優先キュー](#)

[Catalyst 3550 でのウエイテッド ラウンドロビン](#)

[Catalyst 3550 スイッチでの WRED](#)

[Catalyst 3550 スイッチのテール ドロップ](#)

[ギガビット ポートでのキュー サイズの設定](#)

[非ギガビット ポートでのキュー管理とキュー サイズ](#)

[結論](#)

[関連情報](#)

概要

出カスケジューリングは、インターフェイスの出力でオーバーサブスクリプションが過剰に発生した場合に、重要なトラフィックが廃棄されないことを保証します。このドキュメントは、Catalyst 3550 スイッチでの出カスケジューリングに関連するテクニックとアルゴリズムについて包括的に説明しています。また、Catalyst 3550 スイッチにおいて出カスケジューリングの処理を設定および確認する方法についても詳しく説明しています。

前提条件

要件

このドキュメントに関する固有の要件はありません。

使用するコンポーネント

Cisco IOS® ソフトウェア リリース 12.1(12c)EA1 を実行するこの文書に記載されている情報は Catalyst 3550 に基づいています。

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されたものです。このドキュメントで使用するすべてのデバイスは、クリアな（デフォルト）設定で作業を開始しています。ネットワークが稼働中の場合は、コマンドが及ぼす潜在的な影響を十分に理解しておく必要があります。

表記法

ドキュメント表記の詳細は、『[シスコ テクニカル ティップスの表記法](#)』を参照してください。

Catalyst 3550 スイッチでのポートの出力キューイング機能

3550 スイッチのポートには、次の 2 つのタイプがあります。

- ギガビット ポート
- 非ギガビット ポート (10/100 Mbps ポート)

これら 2 つのポートには異なる機能が装備されています。このセクションでは、以降、これらの機能の概要を紹介しています。機能の詳細な説明は、後のセクションで行っています。

ギガビット ポートと非ギガビット ポートの両方がサポートする機能

3550 の各ポートには 4 つの異なる出力キューがあります。これらのいずれかのキューを完全優先キューとして設定できます。残りの各キューは非完全優先キューとして設定され、Weighted Round-Robin (WRR; ウエイトド ラウンドロビン) で処理されます。すべてのポート上で、Class of Service (CoS; サービス クラス) に基づいて、パケットは 4 つの使用可能なキューのいずれかに割り当てられます。

ギガビット ポートのみがサポートする機能

ギガビット ポートは、各キュー内でキュー管理メカニズムもサポートしています。各キューを設定して、Weighted Random Early Detection (WRED; 重み付けランダム早期検出)、または 2 つのしきい値を持つテールドロップを使用できます。また、各キュー (各キューに割り当てられるバッファ) のサイズを調整できます。

非ギガビット ポートのみがサポートする機能

非ギガビット ポートには、WRED や、2 つのしきい値を使用したテールドロップなどのキューイングメカニズムはありません。10/100 Mbps ポートでの FIFO キューイングのみがサポートされます。これらのポートの 4 つの各キューのサイズを変更することはできません。ただし、キューごとに最小 (min) 予約サイズを割り当てられます。

キューに対する CoS のマッピング

このセクションは、3550 がキューへの各パケットの配置をどのように決定するかについて説明しています。パケットは CoS に基づいてキューに配置されます。8 つの使用可能な各 CoS 値は、次の例が示しているキューに対する CoS のマップ インターフェイス コマンドを使用して、4 つの使用可能なキューのいずれかにマッピングされます。

```
(config-if)#wrr-queue cos-map queue-id cos1... cos8
```

次に例を示します。

```
3550(config-if)#wrr-queue cos-map 1 0 13550(config-if)#wrr-queue cos-map 2 2 3 3550(config-if)#wrr-queue cos-map 3 4 53550(config-if)#wrr-queue cos-map 4 6 7
```

この例では次のものが配置されます。

- Q1 に CoS 0 および 1
- Q2 に CoS 2 および 3
- Q3 に CoS 4 および 5
- Q4 に CoS 6 および 7

次のコマンドを発行すると、ポートでのキューに対する CoS のマッピングを確認できます。

```
cat3550#show mls qos interface gigabitethernet0/1 queueing GigabitEthernet0/1...Cos-queue map:cos-qid 0 -  
1 1 - 1 2 - 2 3 - 2 4 - 3 5 - 3 6 - 4 7 - 4...
```

完全優先キュー

完全優先キューは、最初は常に空にされます。そのため、完全優先キューにパケットがあると、そのパケットは即座に転送されます。各パケットが WRR キューのいずれかから転送された後、完全優先キューはチェックされ、必要に応じて空にされます。

完全優先キューは、音声などの遅延/ジッタの影響を受けやすいトラフィック用に特化した設計になっています。完全優先キューは、結果的にその他のキューの処理停滞を発生させる可能性があります。パケットが完全優先キューで待機している場合、その他 3 つの WRR キューに配置されたパケットは転送されません。

ヒント

他のキューの処理停滞を防ぐために、優先キューにどのようなトラフィックが配置されているかに特に注意してください。このキューは、通常、音声トラフィックに使用されます。トラフィックの量は、通常、あまり多くありません。ただし、いずれかのユーザが CoS プライオリティを使用して大量のトラフィック（サイズの大きなファイルの転送やバックアップなど）を完全優先キューに送れるような場合、結果としてその他のトラフィックの処理が停滞します。この問題を避けるために、ネットワークでのトラフィックの分類/アドミッションおよびマーキングに、特別なトラフィックが配置される必要があります。たとえば、次の対策を行うことができます。

- すべての信頼できない送信元ポートには、信頼できないポートの QoS ステータスを使用する。
- Cisco IP Phone ポートには信頼境界機能を使用して、別のアプリケーションの IP Phone 用に設定された信頼状態で Cisco IP Phone ポートが使用されないようにする。
- 完全優先キューが宛先であるトラフィックを監視する。CoS が 5 (Differentiated Services Code Point (DSCP; DiffServ コード ポイント) 46) のポリシングトラフィックの制限をギガビットポート上で 100 MB に設定する。

これらのトピックの詳細は、次のドキュメントを参照してください。

- [Catalyst 3550 のQoS ポリシングおよびマーキングの理解](#)
- 『[QoS の設定](#)』（Catalyst 3500）の「[ポートセキュリティを確保するための信頼境界の設定](#)」セクション

3550 では、1 つのキューを優先キュー（常に Q4）に設定できます。インターフェイスモードで次のコマンドを使用します。

```
3550(config-if)#priority-queue out
```

インターフェイスで優先キューが設定されていない場合、Q4 は標準 WRR キューと見なされます。このドキュメントの「[Catalyst 3550 でのウェイトド ラウンドロビン](#)」セクションは、さらに詳細を説明しています。次と同じ Cisco IOS コマンドを発行すると、インターフェイスに完全優先キューが設定されているかどうかを確認できます。

```
NifNif#show mls qos interface gigabitethernet0/1 queueing GigabitEthernet0/1Egress expedite queue: ena
```

Catalyst 3550 でのウェイトド ラウンドロビン

WRR は、3550 での出力スケジューリングで使用されるメカニズムです。(完全優先キューが存在しない場合) WRR は 3 または 4 つのキューの間で動作します。WRR で使用されるキューはラウンドロビン方式で(送信処理されて)空になり、さらに各キューの重み付けを設定できます。

たとえば、次の表に示すように、キューが別々に処理されるように重み付けを設定できます。

- サーブ WRR 第 1 四半期: 10 %
- サーブ WRR Q2: WRR Q2 の処理 : 時間の 20 %
- サーブ WRR Q3: WRR Q3 の処理 : 時間の 60 %
- サーブ WRR Q4: 10 %

各キューに対して、インターフェイスモードで次のコマンドを発行すると、4 つの重み付け(1 つが各キューに関連付けられている)を設定できます。

```
(config-f)#wrr-queue bandwidth weight1 weight2 weight3 weight4
```

次に例を示します。

```
3550(config)#interface gigabitethernet 0/13550(config-if)#wrr-queue bandwidth 1 2 3 4
```

注重み付けは相対的です。次の値が使用されます。

- $Q1 = \text{重み付け } 1 / (\text{重み付け } 1 + \text{重み付け } 2 + \text{重み付け } 3 + \text{重み付け } 4) = 1 / (1+2+3+4) = 1/10$
- $Q2 = 2/10$
- $Q3 = 3/10$
- $Q4 = 4/10$

WRR は、次の 2 つの方法で実装できます。

- **帯域幅ごとの WRR** : 各重み付けは、送信を許可されている特定の帯域幅を表します。重み付け Q1 は帯域幅の約 10 % を使用することが許可され、Q2 は帯域幅の 20 % を取得する、などのようになります。この方式は、現時点では Catalyst 6500/6000 シリーズのみで実装されています。
- **パケットごとの WRR** : これが 3550 スイッチで実装されているアルゴリズムです。各重み付けは、サイズに関係なく、送信される特定の数のパケットを表します。

3550 はパケットごとに WRR を実装するため、この動作はこのセクションでの設定に次のように適用されます。

- Q1 は 10 パケットのうち 1 パケットを送信する
- Q2 は 10 パケットのうち 2 パケットを送信する
- Q3 は 10 パケットのうち 3 パケットを送信する
- Q4 は 10 パケットのうち 4 パケットを送信する

送信されるパケットはすべて同じサイズであることが考えられます。この段階ではまだ、4 つのキューの間での帯域幅の共有は予測可能です。ところが、キューの間で平均パケットサイズが異なる場合には、輻輳が発生すると、何が送信され何が廃棄されるかには大きな影響が生じます。

たとえば、スイッチには 2 つのフローしか存在しないと仮定します。また仮定として、次の条件が該当するとします。

- CoS が 3 のサイズの小さなインタラクティブ アプリケーション トラフィック (80 バイト [B] フレーム) が 1 Gbps で Q2 に配置されている。
- CoS が 0 のサイズの大きなファイル転送トラフィック (1518-B フレーム) が 1 Gbps で Q1 に配置されている。

スイッチの2つのキューでは1 Gbps でデータが送信される。

両方のストリームは、同じ出力ギガビット ポートを共有する必要がある。Q1 と Q2 の間では等しい重み付けが設定されているとします。WRR はパケットごとに適用され、各キューから送信されるデータの量は、2つのキューの間で異なります。各キューからは同じ数のパケットが送信されますが、スイッチは実際には次の量のデータを送信します。

- Q2 からは 77700 パケット/秒 (pps) = (77700 × 8 × 64) ビット/秒 (bps) (約 52 Mbps)
- Q1 からは 77700 pps = (77700 × 8 × 1500) bps (約 948 Mbps)

ヒント

- ネットワークへの各キューに関する公平なアクセスを実現したい場合は、各パケットの平均サイズを考慮します。各パケットは1つのキューに配置され、重み付けはそれによって変更されることが想定されています。たとえば、各キューが帯域幅の1/4を取得するように、4つの各キューに等しいアクセスを付与する場合、トラフィックは次のようになります。Q1：ベスト エフォートのインターネットトラフィック。トラフィックは256 Bの平均パケットサイズを持っているとします。Q2：パケットが主に1500 Bである、ファイル転送から構成されるバックアップ。Q3：192 Bのパケットで処理されるビデオストリーム。Q4：主に64 Bのパケットから構成されるインタラクティブアプリケーション。これにより次の条件が成立します。Q1はQ4の4倍の帯域幅を消費する。Q2はQ4の24倍の帯域幅を消費する。Q3はQ4の3倍の帯域幅を消費する。
- ネットワークへの等しい帯域幅アクセスを実現するには、次のように設定します。Q1の重み付けを6Q2の重み付けを1Q3の重み付けを8Q4の重み付けを24
- 上記の重み付けを割り当てた場合、輻輳が発生すると、4つのキューの間で等しい帯域幅共有が実現されます。
- 完全優先キューが有効である場合、WRRの重み付けは、残りの3つのキューの間で再分配されます。完全優先キューが有効であり、Q4が設定されていない場合、重み付けが1、2、3、および4の最初の例は次のようになります。Q1 = 1 / (1+2+3) = 6つのうち1パケットQ2 = 6つのうち2パケットQ3 = 6つのうち3パケット次のCisco IOSソフトウェアのshow コマンドを発行すると、キューの重み付けを確認できます。

```
NifNif#show mls qos interface gigabitethernet0/1 queueing GigabitEthernet0/1QoS is disabled. Only one queue is usedWhen QoS is enabled, following settings will be appliedEgress expedite queue: diswrr bandwidth weights:qid-weights 1 - 25 2 - 25 3 - 25 4 - 25
```

緊急優先キューが有効である場合、緊急キューが無効になると、Q4の重み付けのみが使用されます。次に例を示します。

```
NifNif#show mls qos interface gigabitethernet0/1 queueing GigabitEthernet0/1Egress expedite queue: enawrr bandwidth weights:qid-weights 1 - 25 2 - 25 3 - 25 4 - 25 !--- The expedite queue is disabled.
```

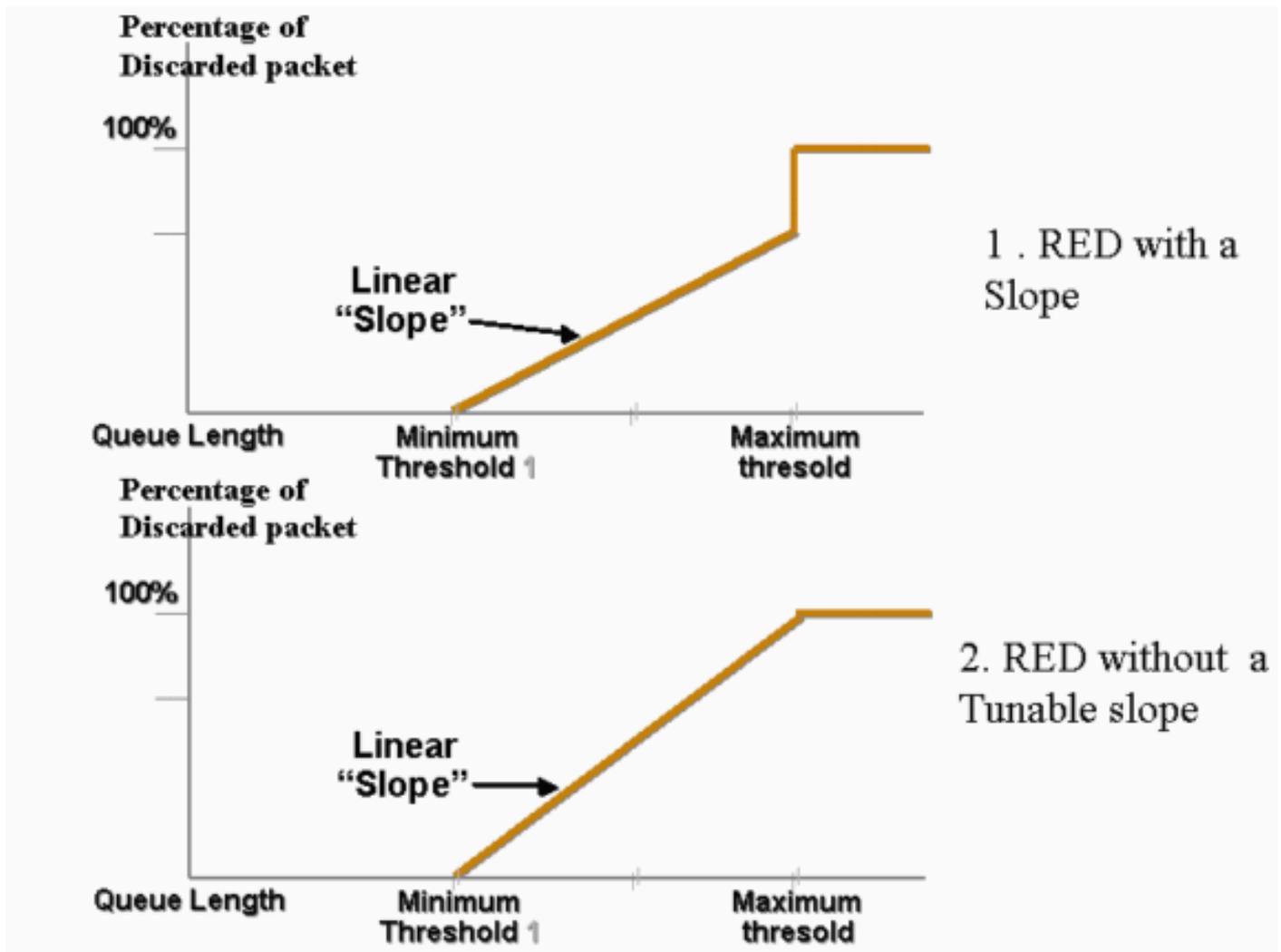
Catalyst 3550 スイッチでの WRED

3550 シリーズ スイッチでは、WRED はギガビット ポートのみで使用できます。WRED は、Random Early Detection (RED; ランダム早期検出) を修正したもので、輻輳回避で使用されます。RED では次のパラメータが定義されています。

- **最小しきい値**：キュー内のしきい値を表します。このしきい値未満では、パケットは廃棄されません。
- **最大 (max) しきい値**：キュー内の別のしきい値を表します。最大しきい値を超えると、すべてのパケットは廃棄されます。
- **スロープ**：最小値と最大値の間でパケットを廃棄する確率。廃棄の確率は、キュー サイズ

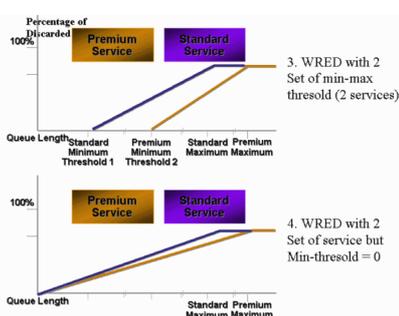
とともに (ある程度のスロープで) 直線的に増加します。
 次のグラフは RED キューでのパケット廃棄の確率を示します。

注RED を実装するすべての Catalyst スイッチでは、スロープを調整できます。



WRED では、異なるサービスが重み付けの対象になります。標準サービスとプレミアム サービスを定義できます。各サービスは、しきい値の別のセットを受け取ります。最小しきい値 1 に到達すると、標準サービスに割り当てられているパケットのみが廃棄されます。最小しきい値 2 に到達すると、プレミアム サービスのパケットのみが廃棄されるようになります。最小しきい値 2 が最小しきい値 1 よりも高い場合、プレミアム サービスのパケットよりも多くの標準サービスのパケットが廃棄されます。次のグラフに、WRED を使用した各サービスの廃棄の確率の例を示します。

注3550 スイッチでは、最大しきいのみを調整でき、最小しきい値を調整できません。最小しきい値は常に 0 にハード設定されています。これにより、現在 3550 で実装されている内容を表す廃棄の確率が決まります。



[3550 スイッチでの WRED](#)」セクションで定義されている同じ DSCP しきい値マップを使用して、DSCP のセットが各テールドロップのしきい値に割り当てられます。しきい値に到達すると、そのしきい値に割り当てられている DSCP を持つすべてのパケットが廃棄されます。次のコマンドを発行すると、テールドロップのしきい値を設定できます。

```
(config-if)#wrr-queue threshold queue-id threshold-percentage1 threshold-percentage2
```

この例では、次のように設定します。

- テールドロップしきい値が 1 の Q1 = 50 % としきい値 2 = 100 %
- しきい値が 1 の Q2 = 70 % としきい値 2 = 100 %

```
Switch(config-if)#wrr-queue threshold 1 50 100Switch(config-if)#wrr-queue threshold 2 70 100Switch(config-if)#wrr-queue threshold 3 60 100Switch(config-if)#wrr-queue threshold 4 80 100
```

[ギガビットポートでのキューサイズの設定](#)

3550 スイッチはセントラルバッファリングを使用します。このことは、ポート単位で固定されたバッファサイズがないことを意味します。ただし、ギガビットポート上にはキューイング可能な固定数のパケットが存在します。この固定数は 4096 です。デフォルトでは、パケットサイズに関係なく、ギガビットポートの各キューは最大 1024 までパケットを持つことができます。ただし、これらの 4096 個のパケットが 4 つのキューの間で分割される方法を変更できます。次のコマンドを発行します。

```
wrr-queue queue-limit Q_size1 Q_size2 Q_size3 Q_size4
```

次に例を示します。

```
3550(config)#interface gigabitEthernet 0/13550(config-if)#wrr-queue queue-limit 4 3 2 1
```

これらのキューサイズのパラメータは相対的です。この例は次のことを示しています。

- Q1 は Q4 の 4 倍の大きさである。
- Q2 は Q4 の 3 倍の大きさである。
- Q3 は Q4 の 2 倍の大きさである。

4096 個のパケットは次の方法で再配布されます。

- $Q1 = [4 / (1+2+3+4)] * 4096 = 1639$ パケット
- $Q2 = 0.3 * 4096 = 1229$ パケット
- $Q3 = 0.2 * 4096 = 819$ パケット
- $Q4 = 0.1 * 4096 = 409$ パケット

次のコマンドを発行すると、4 つのキューの間で分割されたバッファの相対的な重み付けを確認できます。

```
cat3550#show mls qos interface buffersGigabitEthernet0/1Notify Q depth:qid-size1 - 42 - 33 - 24 - 1...
```

また次のコマンドを発行すると、各キューがまだ保持できる空きパケットがいくつあるかを確認できます。

```
(config-if)#show mls qos interface gigabitEthernetx/x statistics WRED drop counts: qid thresh1  
thresh2 FreeQ 1 : 0 0 1639 2 : 0 0 1229 3 : 0 0  
819 4 : 0 0 409
```

FreeQ カウントパラメータはダイナミックです。FreeQ カウンタの値は、最大キューサイズから現在キュー内に存在するパケットの数を引いたものです。たとえば、現在 Q1 に 39 個のパケットが存在する場合、FreeQ カウントでは 1600 個のパケットがフリーです。次に例を示します。

。

```
(config-if)#show mls qos interface gigabitethernetx/x statistics WRED drop counts: qid thresh1
thresh2 FreeQ 1 : 0 0 1600 2 : 0 0 1229 3 : 0 0
819 4 : 0 0 409
```

非ギガビットポートでのキュー管理とキューサイズ

10/100 Mbps ポートで使用できるキュー管理スキームはありません (2つのしきい値を持つ WRED または テール ドロップ はありません)。4 つすべてのキューは FIFO キューです。また、各ギガビットポートに 4096 個のパケットを予約する最大キューサイズも存在しません。10/100 Mbps ポートは、リソース不足によりいっぱいになるまで、各キューにパケットを格納します。キューごとにパケットの最小数を予約できます。この最小値は、デフォルトではキューごとに 100 パケットに設定されています。異なる最小予約値を定義し、いずれかの値を各キューに割り当てた場合、各キューのこの最小予約値を変更できます。

このように変更するには、次の手順を実行します。

1. 各グローバル最小予約値のバッファ サイズを割り当てます。最大 8 つの異なる最小予約値を設定できます。次のコマンドを発行します。

```
(Config)# mls qos min-reserve min-reserve-level min-reserve-buffersize
```

これらの最小予約値はスイッチに対してグローバルです。デフォルトでは、すべての最小予約値が 100 パケットに設定されています。たとえば、150 パケットの最小予約レベル 1 と 50 パケットの最小予約レベル 2 を設定するには、次のコマンドを発行します。

```
nifnif(config)#mls qos min-reserve ?<1-8> Configure min-reserve levelnifnif(config)#mls qos min-reserve 1 ?<10-170> Configure min-reserve buffersnifnif(config)#mls qos min-reserve 1 150nifnif(config)#mls qos min-reserve 2 50
```
2. 各キューにはいずれかの最小予約値を割り当てます。キューに保証されているバッファの数を調べるには、各キューをいずれかの最小予約値に割り当てる必要があります。デフォルトでは、次の条件が適用されます。Q1 は最小予約レベル 1 に割り当てられる。Q2 は最小予約レベル 2 に割り当てられる。Q3 は最小予約レベル 3 に割り当てられる。Q4 は最小予約レベル 4 に割り当てられる。デフォルトでは、すべての最小予約値は 100 です。次のインターフェイスコマンドを発行すると、キューごとに異なる最小予約値を割り当てることができます。

```
(config-if)#wrr-queue min-reserve queue-id min-reserve-level
```

たとえば、最小予約 2 を Q1 に割り当て、最小予約 1 を Q2 に割り当てるには、次のコマンドを発行します。

```
nifnif(config)#interface fastethernet 0/1nifnif(config-if)#wrr-queue min-reserve ?<1-4> queue idnifnif(config-if)#wrr-queue min-reserve 1 ?<1-8> min-reserve levelnifnif(config-if)#wrr-queue min-reserve 1 2nifnif(config-if)#wrr-queue min-reserve 2 1
```

次のコマンドを発行すると、結果の最小予約割り当てを確認できます。

```
nifnif#show mls qos interface fastethernet0/1 buffers
```

```
FastEthernet0/1Minimum reserve buffer size:150 50 100 100 100 100 100 100 !--- This shows the value of all eight min reserve levels.Minimum reserve buffer level select:2 1 3 4 !--- This shows the min reserve level that is assigned to !--- each queue (from Q1 to Q4).
```

結論

3550 上のポートでのキューイングとスケジューリングは次の手順で行います。

1. 各 CoS をいずれかのキューに割り当てる。
2. 必要に応じて、完全優先キューを有効にする。
3. WRR の重み付けを割り当て、キュー内で想定されるパケット サイズを考慮する。
4. キューサイズを変更する (ギガビットポートのみ)。
5. キュー管理メカニズムを有効にする (テールドロップまたは WRED、ギガビットポートのみ)。

適切なキューイングとスケジューリングは、音声/ビデオトラフィックの遅延/ジッタを減らし、

ミッションクリティカルなトラフィックが失われるのを回避できます。スケジューリングのパフォーマンスを最大にするには、必ずこれらのガイドラインに従ってください。

- 信頼または特定のマーキングを使用して、異なるクラスのネットワークに存在するトラフィックを分類する。
- 過度なトラフィックをポリシングする。

関連情報

- [Catalyst 3550 のQoS ポリシングおよびマーキングの理解](#)
- [QoS の設定 - 製品マニュアル](#)
- [LAN 製品に関するサポート ページ](#)
- [LAN スイッチングに関するサポート ページ](#)
- [テクニカルサポートとドキュメント - Cisco Systems](#)