

# Nexus 3500 出カドロップとバッファ QoS

## 内容

[概要](#)

[方法](#)

[出カドロップのチェック](#)

[ドロップがユニキャストかマルチキャストかを確認する](#)

[使用されている出力バッファの判別](#)

[アクティブバッファモニタリングの確認](#)

[カウンタがアクティブに増加](#)

[簡単な出力:](#)

[詳細出力](#)

[しきい値を超えたときにログを生成する](#)

[注目すべき Cisco Bug ID](#)

[よく寄せられる質問 \(FAQ\)](#)

[付録 – 機能情報](#)

[バッファ管理](#)

[スケジューリング](#)

[マルチキャスト低速レシーバ](#)

[実行中のバッファの監視](#)

[ハードウェアの実装](#)

[ソフトウェアの実装](#)

## 概要

このドキュメントでは、Nexus 3500プラットフォームでドロップされるトラフィックのタイプと、このトラフィックがドロップされる出力バッファ(OB)のトラブルシューティングに使用するコマンドについて説明します。

## 方法

1. [出カドロップのチェック](#)
2. [ドロップがユニキャストかマルチキャストかを確認する](#)
3. [使用されている出力バッファの判別](#)
4. [アクティブバッファモニタリングの確認](#)

## 出カドロップのチェック

トラフィックが出力方向にドロップされているかどうかを判別するには、物理インターフェイスの統計情報を確認します。TX方向の「output discard」カウンタが増加するか、またはゼロ以外のカウンタかどうかを確認します。

```

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

```

## ドロップがユニキャストかマルチキャストかを確認する

インターフェイスがトラフィックをドロップしていると判断されたら、**show queuing interface <x/y>** コマンドを入力して、ドロップされたトラフィックがマルチキャストかユニキャストかを確認します。6.0(2)A3(1)より前のリリースでは、出力は次のようになります。

```

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

```

リリース6.0(2)A3(1)以降の出力は次のようになります。

```
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
```

注：マルチキャスト低速レシーバがポートに設定されている場合、機能情報については、ハードウェアの制限により、**show queuing interface Eth<x/y>**コマンドを使用してドロップが追跡されないことを参照してください。Cisco Bug ID [CSCuj21006](#)を参照してください。

## 使用されている出力バッファの判別

Nexus 3500には、出力方向で使用される3つのバッファプールがあります。show hardware internal mtc-usd info port-mappingコマンドの出力には、マッピング情報が表示されます。

```
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.

結果の最初の部分は、OBプール0が45、46、47、48などの前面ポートで使用され、OB1が前面ポート17、18などで使用されていることを示しています。

結果の2番目の部分は、Eth1/1がOB2ポート12にマッピングされ、Eth1/2がOB2ポート13にマッピングされることを示しています。

説明されているポートEth1/7はOB1にマッピングされます。

詳細は、このドキュメントの「[バッファ管理](#)」セクションを参照してください。

## アクティブバッファモニタリングの確認

この機能の詳細については、[Cisco Nexus 3548のアクティブなバッファモニタリングに関するホワイトペーパー](#)と、このドキュメントのセクションを参照してください。



```

0      0
09/30/2013 19:46:56      0      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      0      251
0      0
09/30/2013 19:46:54      0      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      0      250
0      0
09/30/2013 19:46:52      0      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      0      249
0      0

```

...

各行の情報は、2番目の間隔で記録されます。各列は、バッファの使用状況を表します。コマンドの結果で説明したように、列「384」に対してレポートされたゼロ以外の値がある場合は、ABMがOB使用率をポーリングしたときに、バッファ使用率が0 ~ 384 KBであったことを意味します。0以外の数値は、使用状況が報告された回数です。

これらの結果は、Eth1/7の最後の10秒間に、OB1が每秒249 ~ 253回の平均使用率が5.376 MBであることを示しています。このトラフィックのバッファをクリアするには、4298マイクロ秒(us)がかかります。

## しきい値を超えたときにログを生成する

ドロップカウンタとバッファの使用率が定期的増加する場合は、しきい値を設定し、しきい値を超えたときにログメッセージを生成できます。

```

logging level mtc-usd 5
hardware profile buffer monitor unicast sampling 10 threshold 4608

```

このコマンドは、10ナノ秒の間隔でユニキャストトラフィックを監視するように設定され、バッファの75%を超えるとログを生成します。

スケジューラを作成して、ABM統計情報とインターフェイスカウンタの出力を1時間ごとに収集し、ブートフラッシュファイルに追加することもできます。次に、マルチキャストトラフィックの例を示します。

```

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

```

## 注目すべき Cisco Bug ID

- Cisco Bug ID [CSCum21350](#):高速ポートフラップにより、同じQoSバッファ内のすべてのポートで、すべてのTXマルチキャスト/ブロードキャストトラフィックが廃棄されます。これは、

リリース6.0(2)A1(1d)以降で修正されています。

- Cisco Bug ID [CSCuq96923](#):マルチキャストバッファブロックがスタックしているため、出力マルチキャスト/ブロードキャストのドロップが発生します。この問題はまだ調査中です。
- Cisco Bug ID [CSCva20344](#):Nexus 3500バッファブロック/ロックアップ : TXマルチキャストまたはブロードキャストなし再現不可能な問題。リリース6.0(2)U6(7)、6.0(2)A6(8)、および6.0(2)A8(3)で修正される可能性があります。
- Cisco Bug ID [CSCvi93997](#): Cisco Nexus 3500スイッチの出力バッファブロックスタック。これは、リリース7.0(3)I7(8)および9.3(3)で修正されています。

## よく寄せられる質問 ( FAQ )

ABMはパフォーマンスまたは遅延に影響しますか。

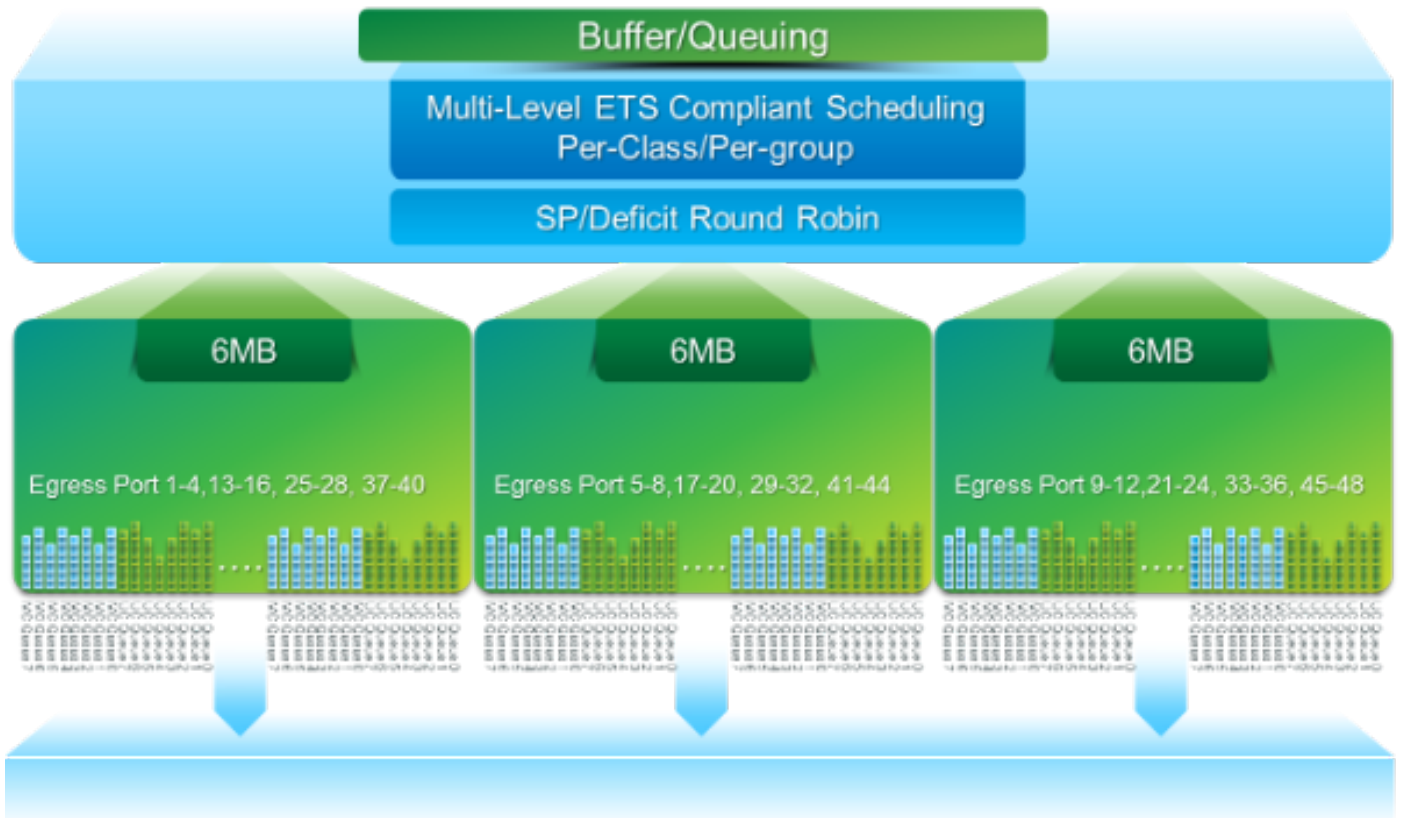
いいえ。この機能はデバイスの遅延やパフォーマンスに影響を与えません。

低いABMハードウェアポーリング間隔の影響は何ですか。

デフォルトでは、ハードウェアポーリング間隔は4ミリ秒です。この値は10ナノ秒まで低く設定できます。ハードウェアポーリング間隔が短いため、パフォーマンスや遅延の影響はありません。ソフトウェアが1秒ごとにポーリングする前にヒストグラムカウンタをオーバーフローしないように、デフォルトの4ミリ秒のハードウェアポーリングが選択されています。ハードウェアポーリング間隔を小さくすると、255サンプルのハードウェアカウンタが飽和状態になる可能性があります。デバイスは、CPUおよびメモリの制限により低いハードウェアポーリングに一致させるために、1秒未満のソフトウェアポーリングを処理できません。このホワイトペーパーには、ハードウェアのポーリング間隔が短い例と、その使用例が記載されています。

## 付録 – 機能情報

### バッファ管理

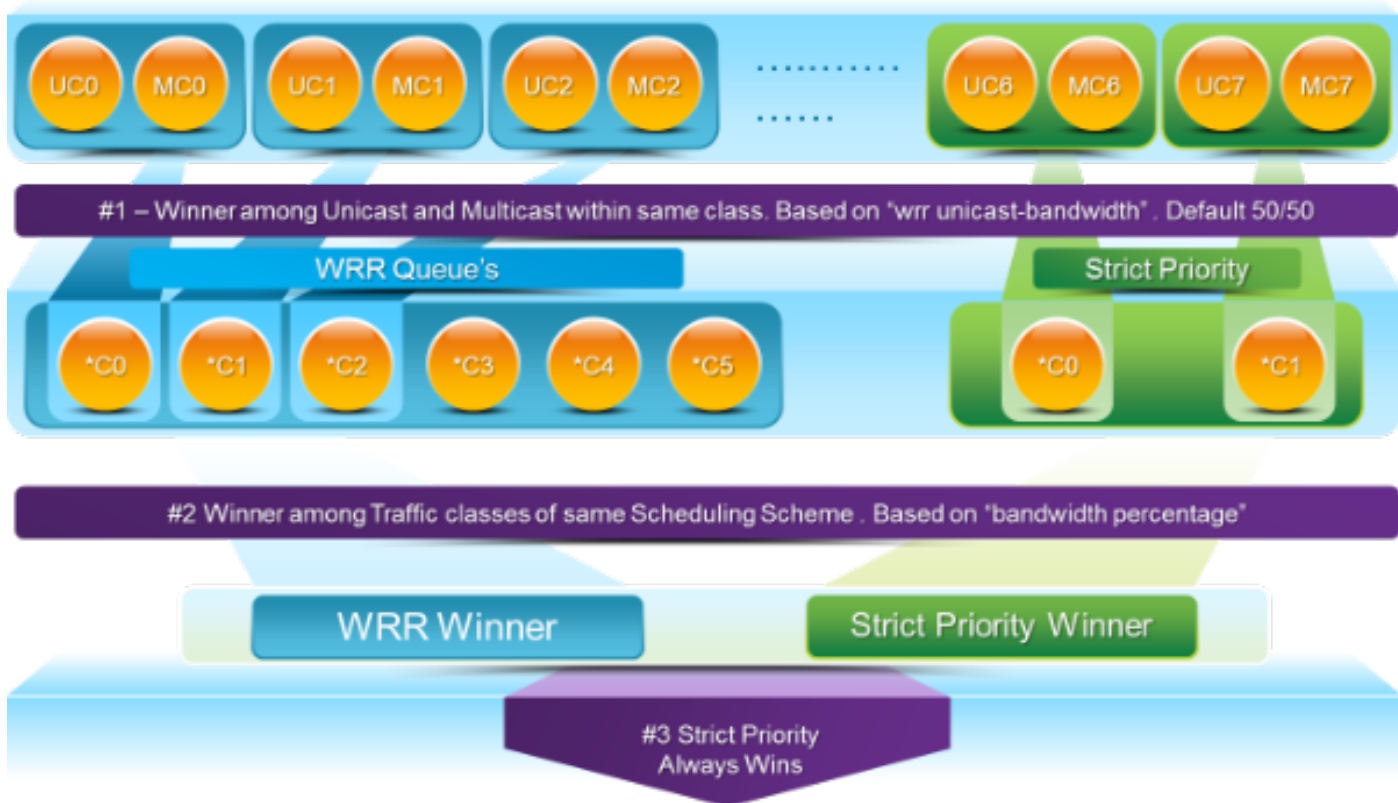


- 3つのOBブロックで共有される18 MBパケットバッファ：最大4 MB予約済み：設定された最大伝送ユニット(MTU)に基づくサイズ (ポート単位の合計は2 x MTUサイズx有効なQoSグループの数) 最大14 MBの共有：合計バッファの残り最大767 KBのOB:CPU宛ての packets の場合は0
- 各OBの6 MBは、16ポートのセットで共有されます(show hardware internal mtc-usrd info port-mappingコマンド)。

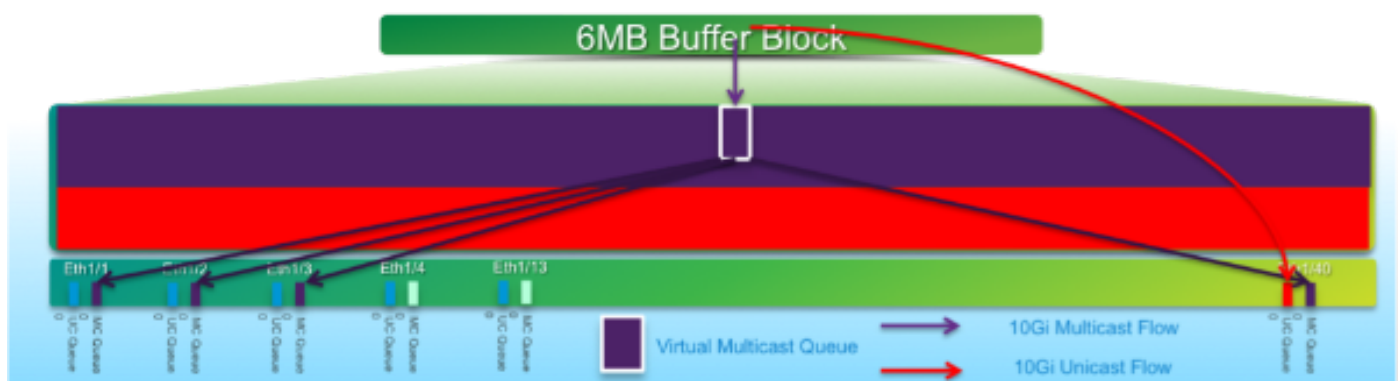
## スケジューリング

3層スケジューリング：

- 基本アーキテクチャ概要
- 同じスケジューリング方式のトラフィッククラス
- スキーム全体のトラフィッククラス



## マルチキャスト低速レシーバ



このダイアグラムでは、

- 持続的な輻輳は、1 G Eth1/40で発生します。
- バッファブロックのその他のマルチキャストレシーバ(Eth1/1 ~ 3)は、マルチキャストスケジューリングの動作により影響を受けます。他のバッファブロックのレシーバは影響を受けません。
- 「マルチキャスト低速レシーバ」をe1/40に適用すると、輻輳していないポートでのトラフィック損失を回避できます。
- 「マルチキャスト低速レシーバ」は、Eth1/40でマルチキャストを10 Gレートで排出します。輻輳したポートでは引き続き廃棄が発生する可能性があります。
- hardware profile multicast slow-receiver port <x>コマンドで設定されます。

## 実行中のバッファの監視

この機能の概要については、『[Cisco Nexus 3548アクティブバッファモニタリング](#)』ホワイトペーパーを参照してください。



## ハードウェアの実装

- ASICには18のバケットがあり、各バケットはバッファ使用率の範囲 ( 0 ~ 384KB、385 ~ 768KBなど ) に対応します。
- ASICは、4ミリ秒 ( デフォルト ) ごとにすべてのポートのバッファ使用率をポーリングします。このASICポーリング間隔は、10ナノ秒まで設定可能です。
- 各ハードウェアポーリング間隔のバッファ使用率に基づいて、対応する範囲のバケットカウンタが増分されます。つまり、ポート25が500 KBのバッファを消費すると、バケット#2(385 ~ 768KB)カウンタが増分されます。
- このバッファ使用率カウンタは、ヒストグラム形式でインターフェイスごとに維持されます。
- 各バケットは8ビットで表されるため、カウンタは255でカウントアウトし、ソフトウェアがデータを読み取るとリセットされます。

## ソフトウェアの実装

- ソフトウェアは1秒ごとにASICをポーリングして、すべてのヒストグラムカウンタをダウンロードしてクリアします。
- これらのヒストグラムカウンタは、1秒の精度で60分間メモリに保持されます。
- また、バッファヒストグラムを1時間ごとにブートフラッシュにコピーし、さらに分析するためにアナライザにコピーすることもできます。
- これにより、すべてのポートに対して2時間分のバッファヒストグラムデータが維持されます。これは、メモリ内で最新の1時間、ブートフラッシュ内で2時間目です。