CHAPTER 3

レイヤ 2 スイッチングの問題のトラブルシューティング

レイヤ 2 は、コンピュータ ネットワーキングの Open Systems Interconnection（OSI）モデルのデータ リンク層です。

この章では、Cisco Nexus 5000 シリーズ スイッチのレイヤ 2 スイッチングで起こり得る問題を特定し、解決する方法について説明します。

この章で説明する内容は、次のとおりです。

MAC アドレス テーブル

データ トラフィックのフラッディング

データが転送されず、代わりに VLAN のすべてのポートにフラッディングされます。

考えられる原因

ループの検出が原因で、MAC アドレスの学習がディセーブルになっています。この場合は、重大度 2 の Syslog メッセージ、STM_LOOP_DETECT が受信されています。

解決方法

180 秒待機した後、学習が自動的にイネーブルになります。重大度 2 の Syslog メッセージ、STM_LEARNING_RE_ENABLE が受信されます。

考えられる原因

MAC アドレス テーブルがいっぱいになっています。この場合は、重大度 2 の Syslog メッセージ、STM_LIMIT_REACHED が受信されています。

解決方法

180 秒待機した後、MAC テーブルがフラッシュされ、学習が自動的にイネーブルになります。あるいは、一部の MAC エントリの期限が切れて学習済みエントリの総数が 1500 より少なくなるまで待つか、「clear mac address-table dynamic [address <mac>]」コマンドを実行してエントリをクリアします。こうすると、新しい MAC エントリを学習するための空き領域が作成されます。重大度 2 の Syslog メッセージ、STM_LEARNING_RE_ENABLE が受信されます。

考えられる原因

学習の過負荷が原因で（つまり、新しいアドレスが短時間であまりにも多すぎたため）、MAC アドレスの学習がディセーブルになっています。この場合は、重大度 4 の Syslog メッセージ、STM_LEARNING_OVERLOAD が受信されています。

解決方法

120 秒待機した後、学習が自動的にイネーブルになります。

MAC アドレスが学習されない

スイッチによって MAC アドレスが学習されません。これが起こると、MAC アドレスは MAC テーブルに登録されません。

考えられる原因

ループの検出が原因で、MAC アドレスの学習がディセーブルになっています。この場合は、重大度 2 の Syslog メッセージ、STM_LOOP_DETECT が受信されています。

解決方法

180 秒待機した後、学習が自動的にイネーブルになります。重大度 2 の Syslog メッセージ、STM_LEARNING_RE_ENABLE が受信されます。

考えられる原因

MAC アドレス テーブルがいっぱいになっています。この場合は、重大度 2 の Syslog メッセージ、STM_LIMIT_REACHED が受信されています。

解決方法

180 秒待機した後、MAC テーブルがフラッシュされ、学習が自動的にイネーブルになります。あるいは、一部の MAC エントリの期限が切れて学習済みエントリの総数が 1500 より少なくなるまで待つか、「clear mac address-table dynamic [address <mac>]」コマンドを実行してエントリをクリアします。こうすると、新しい MAC エントリを学習するための空き領域が作成されます。重大度 2 の Syslog メッセージ、STM_LEARNING_RE_ENABLE が受信されます。

考えられる原因

学習の過負荷が原因で（つまり、新しいアドレスが短時間であまりにも多すぎたため）、MAC アドレスの学習がディセーブルになっています。この場合は、重大度 4 の Syslog メッセージ、STM_LEARNING_OVERLOAD が受信されています。

解決方法

120 秒待機した後、学習が自動的にイネーブルになります。

考えられる原因

着信データ トラフィック用の出力パスが設定されていません。スイッチから発信されるデータのパスがない場合、そのデータ ストリームから MAC アドレスは学習されません。

解決方法

データの送出パスを設定します。

たとえば、データが着信するインターフェイスを除くすべてのインターフェイスで VLAN がイネーブルになっていない場合があります。あるいは、送出インターフェイスがダウンしている場合もあります。この場合は、それらのインターフェイスをアップする必要があります。

VPC セットアップでのトラフィック フラッディング

VPC シナリオの下でデータが転送されず、代わりにフラッディングされます。

考えられる原因

MAC アドレスが 1 台のスイッチのみで学習されています。通常、この状況は VPC ピアとの MAC アドレスの同期に関するバグです。

解決方法

MAC アドレスが学習されたスイッチから MAC アドレスをクリアします。これにより、MAC アドレスの新しい学習と VPC スイッチ間での MAC アドレスの同期が行われます。

スパニング ツリー プロトコル

メッセージ「BPDUGuard errDisable」で HIF がダウンする

メッセージ「BPDUGuard errDisable」に伴って HIF がダウンします。

考えられる原因

デフォルトでは、HIF は BPDU ガードがイネーブルになった STP エッジ モードになります。つまり、HIF はホストまたは非スイッチング デバイスに接続するよう想定されています。HIF が BPDU を送信する非ホスト デバイスまたはスイッチに接続している場合、その HIF は BPDU の受信時にエラー ディセーブルになります。

解決方法

HIF およびピア接続デバイスで BPDUfilter をイネーブルにします。このフィルタをイネーブルにすると、HIF は BPDU を送信または受信しません。次のコマンドを使用して、ポートの STP ポート ステートの詳細を確認します。

- 「show spanning-tree interface <id> detail」
- 「show spanning-tree interface <id>」

スイッチで FWM-2-STM_LOOP_DETECT が検出され、動的学習がディセーブルになる

スイッチで FWM-2-STM_LOOP_DETECT が検出されると、動的学習がディセーブルになります。

考えられる原因

- 不適切な STP ポート ステートのコンバージェンスが原因で、MAC アドレスが移動しています。
- STP ステートがコンバージされて正しい状態にあるときにデータの送信元がすべてのスイッチを物理的に横断していることが原因で、MAC アドレスが移動しています。

次のコマンドを使用して、スイッチの VLAN 上の STP ポート ステートを確認します。

- 「show spanning-tree」
- 「show spanning-tree vlan <id>」

解決方法

- 正しい STP コンバージェンスを確認し、関係図内のすべてのスイッチで STP ポートステートをチェックします。また、競合がないこと、および不適切なポートステートがないことも確認します。
- 物理的に移動しているデータ フレームの送信元が特定された場合は、高速な連続的移動を停止するよう送信元を制御します。
- デフォルトでは、動的学習は 180 秒後に再開します。その時点で、すべての STP 競合または不整合は解決されます。

STP ブロッキング ステート「BLK*(Type_Inc)」でポートがスタックしている

STP ブロッキング ステート「BLK*(Type_Inc)」でポートがスタックしています。

考えられる原因

アクセス ポートが相手側のトランク ポートに接続するときに、アクセス ポートでタイプが一致していない可能性があります。リンクに不適切な設定があることを示すため、そのポートは BLK*(Type_Inc) になります。次のコマンドを使用して、ポートの STP ポート ステートの詳細を確認します。

- 「show spanning-tree interface <id> detail」
- 「show spanning-tree interface <id>」

解決方法

リンクの両端（ポート）で設定されている switchport モードを確認します。両者が同じモードになるようにします。両方をアクセス モードまたはトランク モードにする必要があります。モードが同期したら、ポートは不一致状態から正常な状態に移行します。

STP ブロッキング ステート「BLK*(PVID_Inc)」でポートがスタックしている

STP ブロッキング ステート「BLK*(PVID_Inc)」でポートがスタックしています。

考えられる原因

トランク リンク上でネイティブ VLAN の不一致があるときに、PVID が一致していない可能性があります。これが起こると、ポート ステートは「BLK* (PVID_Inc)」になります。次のコマンドを使用して、ポートの STP ポート ステートの詳細を確認します。

- 「show spanning-tree interface <id> detail」
- 「show spanning-tree interface <id>」

解決方法

リンクの両端（ポート）で設定されているネイティブ VLAN を確認します。両者が同じネイティブ VLAN を持つようにします。ネイティブ VLAN が同期したら、ポートは不一致状態から正常な状態に移行します。

STP ブロッキング ステート「BLK*(Loop_Inc)」でポートがスタックしている

STP ブロッキング ステート「BLK*(Loop_Inc)」でポートがスタックしています。

考えられる原因

この状況は、ポートでループガードが設定されていて、ポートで BPDU の受信が停止したときに起こります。これは単方向リンク障害が発生したときにループを防止する働きがあります。ただし、そのポートは「BLK* (Loop_Inc)」ステートになります。次のコマンドを使用して、ポートの STP ポート ステートの詳細を確認します。

- 「show spanning-tree interface <id> detail」
- 「show spanning-tree interface <id>」

解決方法

リンクの両端（ポート）で設定されているネイティブ VLAN を確認します。両者が同じネイティブ VLAN を持つようにします。ネイティブ VLAN が同期したら、ポートは不一致状態から正常な状態に移行します。

マルチキャスト

IGMP 加入の送信元 MAC アドレスが学習される

この状況では、IGMP 加入の送信元 MAC アドレスが学習されます。しかし、MAC アドレス空間を節約するため、スイッチでは通常、IGMP 加入の送信元 MAC アドレスは学習されません。

考えられる原因

加入の受信と ISSU の実行が同時に行われるとこの状況が起こる場合があります。

解決方法

加入が停止した場合、MAC アドレスは期限切れになります。あるいは、「clear mac address-table dynamic mac <mac>」コマンドを使用して明示的に MAC アドレスをクリアすることもできます。

マルチキャスト データ トラフィックがホストで受信されない

ホストでマルチキャスト データ トラフィックが受信されません。

考えられる原因

マルチキャストへの加入が登録されていません。

解決方法

- ホスト アプリケーションから加入が送信されていることを確認します。
- 「show vlan id <vlan>」コマンドを使用して、加入が送信されている VLAN にスイッチ ポートが設定されているかどうかを確認します。
- 「show vlan id <vlan>」コマンドを使用して、該当する VLAN がアクティブかどうかを確認します。
- 「show spanning-tree vlan <vlan>」コマンドを使用して、スイッチ ポートが STP フォワーディング ステートにあるかどうかを確認します。

ホストがグループに登録されているのにマルチキャスト データ トラフィックが受信されない

ホストがグループに登録されているのにマルチキャスト データ トラフィックが受信されません。

考えられる原因

IGMP プロセスと FWM プロセス間の通信にバグがある可能性があります。

次のコマンドの出力を調べます。

- 「show ip igmp snooping groups vlan 1001」
- 「show mac address-table multicast vlan 1001 igmp-snooping」
- 「show platform fwm info vlan 1001 all_macgs verbose」

解決方法

ホスト インターフェイスで「shut/no-shut」操作を実行し、加入をもう一度送信します。

VPC セットアップでマルチキャスト トラフィックがフラッディングする

VPC セットアップでマルチキャスト トラフィックがフラッディングします。

考えられる原因

いずれかのスイッチで IGMP スヌーピングがディセーブルになっています。

解決方法

両方のスイッチで IGMP スヌーピングをイネーブルにします。

（注）　リンク ローカル IP アドレス（つまり、224.0.0.X）のグループは作成されません。

VLAN

Nexus 5000 に VTP サーバを実行しているスイッチと同じ VLAN がない

Nexus 5000 の VLAN が、VTP サーバを実行しているスイッチの VLAN と同じではありません。

考えられる原因

Nexus 5000 では現在、透過モードの VTP のみがサポートされています（4.2(1)N1(1) 以降のリリース）。

解決方法

この状況は、VLAN をローカルに設定する必要があることを示します。ただし、次のコマンドを使用すると、VTP クライアントとサーバが Nexus 5000 を通じて通信できるようになります。

```
switch(config)# feature vtp
switch(config)# vtp mode transparent
switch(config)# exit
switch# show vtp status
```

VLAN が作成できない

VLAN が作成できません。

考えられる原因

内部 VLAN 範囲が使用されています。

解決方法

内部使用のために予約されていない VLAN 番号を使用します。

（注）　3968 ～ 4047 の VLAN 番号は内部使用のために予約されています。

例：

```
switch(config)# vlan ?
  <1-3967,4048-4093>  VLAN ID 1-4094 or range(s): 1-5, 10 or 2-5,7-19
```

インターフェイス VLAN がダウンしている

インターフェイス VLAN がダウンしています。

考えられる原因

VLAN が作成されていません。

解決方法

VLAN <###> がまだ作成されていなくても、NX-OS では「interface vlan <###>」の設定が許可されます。その結果、「interface vlan <###>」はアップしません。「show vlan」コマンドを使用して、VLAN <###> が存在するかどうかを確認します。存在しない場合は、「vlan <###>」コマンドを使用して VLAN を作成します。VLAN を作成した後、アップするためにインターフェイス VLAN をバウンスする必要があります。

例：

```
switch(config)# int vlan 600
switch(config-if)# no shut
switch(config-if)# sh int vlan 600 brief

--------------------------------------------------------------------------------
Interface Secondary VLAN(Type)                    Status Reason
--------------------------------------------------------------------------------
Vlan600   --                                      down   other
switch(config)# show vlan id 600
VLAN 600 not found in current VLAN database
switch(config-if)# vlan 600
switch(config)# show vlan id 600

VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
600  VLAN0600                         active    Po1, Po11, Po30, Po31
switch(config-if)# int vlan 600
switch(config-if)# shut
switch(config-if)# no shut
switch(config-if)# show int vlan 600 brief

--------------------------------------------------------------------------------
Interface Secondary VLAN(Type)                    Status Reason
```

```
--------------------------------------------------------------------------------
Vlan600   --                                      up      --
```

ポートにアクセスするようにインターフェイスを設定しても VLAN <###> を通過できない

VLAN <###> を許可するためにインターフェイスをポートにアクセスするよう設定しても、VLAN <###> を通過できません。

考えられる原因

VLAN が作成されていません。

解決方法

NX-OS では、インターフェイスに対して「switchport access vlan <###>」コマンドを実行しても、VLAN <###> は自動的に作成されません。「vlan <###>」コマンドを使用して VLAN <###> を明示的に作成する必要があります。「show vlan」コマンドを使用して、VLAN <###> が存在するかどうかを確認します。存在しない場合は、「vlan <###>」コマンドを使用して VLAN を作成します。

VLAN が作成できない

VLAN が作成できません。

考えられる原因

VLAN リソースがすべて使用されています。

解決方法

Nexus 5000 では、アクティブな VLAN/VSAN の最大数はスイッチあたり 512 です（VSAN 用に 31、残りは VLAN 用）。「show resource vlan」コマンドを使用して、使用可能な VLAN の数を確認します。

例：

```
switch(config)# show resource vlan

     Resource              Min       Max      Used    Unused     Avail
    -----------           -----     -----    ------  --------   -------
     vlan                 16        512        25         0       487
```

SVI が作成できない

SVI が作成できません。

考えられる原因

「interface-vlan」機能がイネーブルになっていません。

解決方法

SVI を設定する前に、「interface-vlan」機能をイネーブルにする必要があります。「show feature」コマンドを使用して、イネーブルになっている機能を確認します。

例：

```
switch(config)# feature interface-vlan
switch(config)# show feature
```

```
Feature Name          Instance  State
--------------------  --------  --------
tacacs                1         disabled
lacp                  1         enabled
interface-vlan        1         enabled
private-vlan          1         enabled
udld                  1         enabled
vpc                   1         enabled
fcoe                  1         disabled
fex                   1         enabled
```

プライベート VLAN（PVLAN）が作成できない

プライベート VLAN（PVLAN）が作成できません。

考えられる原因

「private-vlan」機能がイネーブルになっていません。

解決方法

PVLAN を設定する前に、「private-vlan」機能をイネーブルにする必要があります。そうすると、PVLAN コマンドが使用可能になります。「show feature」コマンドを使用して、イネーブルになっている機能を確認します。

例：

```
switch(config)# feature private-vlan
switch(config)# show feature
Feature Name          Instance  State
--------------------  --------  --------
tacacs                1         disabled
lacp                  1         enabled
interface-vlan        1         enabled
private-vlan          1         enabled
udld                  1         enabled
vpc                   1         enabled
fcoe                  1         disabled
fex                   1         enabled
```

レジスタとカウンタ

ドロップの識別

Nexus 5000 でフレームがドロップされるときは論理的かつ物理的な原因があります。また、スイッチ アーキテクチャのカットスルー特性のためにフレームをドロップできない場合もあります。ドロップする必要があるフレームがカットスルー パスで切り替えられている場合、唯一のオプションはイーサネット Frame Check Sequence（FCS; フレーム チェック シーケンス）をストンプすることです。フレームをストンプするには、CRC チェックを通過しない既知の値に FCS を設定します。こうすると、このフレームのパスの後の方で後続の CRC チェックが失敗します。これにより、ダウンストリームのストアアンドフォワード デバイスまたはホストがこのフレームをドロップできます。

（注）　フレームが 10 Gb/秒インターフェイスで受信されたとき、そのフレームはカットスルー パスにあるものと見なされます。

次の出力例は、特定のインターフェイスで見られたすべての廃棄とドロップ（キューイング ドロップを除く）を示します。キューイング ドロップは想定どおりの動作である場合があり、エラーの結果生じることもあります（ドロップは廃棄よりも一般的です）。

例：

```
switch# show platform fwm info pif ethernet 1/1 ...
Eth1/1 pd: tx stats: bytes 19765995 frames 213263 discard 0 drop 0
Eth1/1 pd: rx stats: bytes 388957 frames 4232 discard 0 drop 126
```

一部のコマンドでは、ポートが存在するチップを知っておく必要があります。

次の例では、チップの名称は「Gatos」です。この例は、どの Gatos とどの Gatos ポートがイーサネット 1/1 に関連付けられているかを示しています。

```
switch# show hardware internal gatos port ethernet 1/1 | include
instance|mac
         gatos instance     : 7 <- Gatos 7
         mac port           : 2 <- Port 2
         fw_instance        : 2
```

想定されたドロップ/論理的ドロップ

通常運用時、Nexus 5000 は論理的帰結に基づいて転送できないフレームに遭遇します。

たとえば、特定のインターフェイスで MAC アドレスを学習し、そのインターフェイスで受信したトラフィックが宛先としてその送信元インターフェイス上の MAC アドレスを持つ場合、このフレームは転送できません。これは既知のアドレスであり、したがってフラッディングできず、着信インターフェイスからトラフィックは送出できません。これはレイヤ 2 トポロジのループを回避するための要件です。

入力ポートが VLAN 内の唯一のポートである場合は、次の例に示すエラー カウンタが増加します。

例（前の例と同じ Gatos インスタンス）：

```
switch# show platform fwm info gatos-errors 7
Printing non zero Gatos error registers:
DROP_SRC_MASK_TO_NULL     9
```

（注）　「show platform fwm info gatos-errors」コマンドは、ある特定のドロップについてカウンタを 3 回増加させます。

その他の想定されたドロップ

ドロップ	説明
VLAN_MASK_TO_NULL	CPU インターフェイス宛てのトラフィック。 実際には VLAN ではありません。

ドロップ	説明
DROP_NO_FABRIC_SELECTED	VLAN_MASK_TO_NULL とともに増加します。
DROP_INGRESS_ACL	アクセスリストがフレームと一致した場合に増加します。 ACL が適用されていない場合には、NX-OS を DoS 攻撃から守るためにハードウェアでレート制限がイネーブルにされている場合に大量の CPU 宛てトラフィックが受信されると、このカウンタが増加します。

キューがいっぱいになった

キューがいっぱいになったときは、入力インターフェイス上の個々のキューで廃棄を増やす必要があります。

例：

```
switch# show queuing interface e1/1
Ethernet1/1 queuing information:
   TX Queuing
     qos-group  sched-type  oper-bandwidth
         0       WRR             50
         1       WRR             50

   RX Queuing
     qos-group 0
     q-size: 243200, HW MTU: 1600 (1500 configured)
     drop-type: drop, xon: 0, xoff: 1520
     Statistics:
         Pkts received over the port             : 0
         Ucast pkts sent to the cross-bar        : 0
         Mcast pkts sent to the cross-bar        : 0
         Ucast pkts received from the cross-bar  : 0
         Pkts sent to the port                   : 0
     <b>    Pkts discarded on ingress               : 0 </b>
         Per-priority-pause status               : Rx (Inactive), Tx
(Inactive)

     qos-group 1
     q-size: 76800, HW MTU: 2240 (2158 configured)
     drop-type: no-drop, xon: 128, xoff: 240
     Statistics:
         Pkts received over the port             : 0
         Ucast pkts sent to the cross-bar        : 0
         Mcast pkts sent to the cross-bar        : 0
         Ucast pkts received from the cross-bar  : 0
         Pkts sent to the port                   : 0
     <b>   Pkts discarded on ingress               : 0 </b>
         Per-priority-pause status               : Rx (Inactive), Tx
(Inactive)

   Total Multicast crossbar statistics:
     Mcast pkts received from the cross-bar      : 0
```

MTU 違反

Nexus 5000 は 10 Gb/秒ではカットスルー スイッチです。これは、MTU のチェックはできるものの、長さがわかるときにはすでにフレームの送信が開始されていることを意味します。したがって、フレームはドロップできません。このようなフレームは MTU に達した後に切り捨てられ、CRC 値がストンプされます。入力インターフェイスでは Rx Jumbo が増加し、出力インターフェイスでは Tx CRC と Tx Jumbo が増加します。

- 「show interface」コマンドまたは「show hardware internal gatos port e1/1 counters rx」コマンドでジャンボ フレームが示される場合、これはそれらのフレームがドロップされたことを意味するものではありません。ジャンボ フレームとは単に、受信または送信された 1500 バイトを超えるイーサネット フレームのことです。
- 「show queuing interface ex/y」コマンドは、現在の（クラスごとの）MTU の設定値を示します。
- MTU 違反に起因するドロップを確認するには、「show hardware internal gatos counters interrupt match mtu*」コマンドを使用します。
- 「show hardware internal gatos port ethernet 1/1 | include instance|mac」コマンドによって出力される、Gatos 番号と fw_instance に一致するカウンタは、MTU 違反が発生してフレームがストンプされたことを示す指標です。

CRC エラーの処理

カットスルー ポート上の FCS で CRC エラーが発生した場合は、「show interface」の Rx CRC カウンタが増加します。ただし、FCS はワイヤ上のイーサネット フレームの最後にあるため、CRC エラーのフレームはドロップできません。

出力インターフェイスの Tx CRC エラーが増加し、パス上の次のデバイスに伝播します。

Nexus 5000 が CRC を伝播または生成しているかどうかを確認するには、「show hardware internal gatos counters interrupt match stomp」コマンドを使用します。

- ストンプ値が存在する場合は、そのインターフェイス上に一致する CRC 値があります。
- Rx CRC 値が存在する場合は、すでにエラーのある状態でスイッチポートに入ってきたことがわかります。接続されているデバイスに移動し、エラーをさかのぼって追跡できます。