VXLAN パスの検証と検証

VXLAN OAM またはVXLAN NGOAM

VXLAN Operations, Administration, and Maintenance（OAM）または Next Generation Operations, Administration, and Maintenance（NGOAM）は、

- インストール、モニタリング、およびトラブルシューティング中の VXLAN オーバーレイ ネットワークの管理を強化
- VXLAN ネットワークの問題を診断するために、ping、traceroute、または pathtrace に似たトラブルシューティング ツール（VXLAN OAM ツール）を提供します。

これらのプロトコルは、診断チャネルを使用して宛先を追跡し、重要な情報を伝送します。

VXLAN **OAM** は、**NGOAM** とも呼ばれます。

VXLAN OAM ツール

VXLAN OAM ツールは、次の表に示すように分類されます。OAM ツールの詳細については、障害分離と検証ツールの例 （5 ページ） を参照してください。

カテゴリ	ツール
障害検査	loopback メッセージ
障害の隔離	Pathtrace メッセージ

VXLAN OAM ペイロード

OAM チャネルは、これらの OAM パケットに存在する VXLAN ペイロードのタイプを識別するために使用されます。次の種類のペイロードがサポートされています。

- **従来のICMPチャネル**：これらのチャネルは、新しい OAM パケット形式をサポートしていない従来のホストまたはスイッチとの通信を容易にするために使用されます。
- **NVO3 ドラフト Tissa チャネル**：これらのチャネルは、サポートされているホストまたはスイッチとの通信を容易にするために使用され、重要な診断情報を伝送します。「チャネル」という用語は、データを配信するためのメカニズムまたは経路を示しています。

  VXLAN NVO3 ドラフトの Tissa OAM メッセージは、次を使用して識別されます。

  - **予約済み OAM EtherType**：これは OAM パケットを認識するために使用される特定の識別子です。Ethertype は、どのプロトコルがイーサネット フレームのペイロードにカプセル化されるかを示すために使用されます。
  - **予約済み送信元 MAC アドレス**：または、既知の予約済み元 MAC アドレスを使用して OAM パケットを識別することもできます。MAC アドレスは、物理ネットワーク セグメント上での通信のためにネットワーク インターフェイスに割り当てられる固有の識別子です。

VXLAN NGOAM の注意事項と制約事項

VXLAN NGOAM には、次の注意事項と制約事項があります。

Cisco NX-OS リリース 10.2(3)F 以降、中間ノードで NGOAM 機能を使用する **feature nv overlay** コマンドを使用して VXLAN 機能を有効にする必要はありません。

VXLAN NGOAM でサポートされるプラットフォームとリリース

サポートされるリリース	サポートされるプラットフォーム
9.3(3) 以降	Cisco Nexus 9300-FX/FX2/GX シリーズスイッチ
9.3(5) 以降	Cisco Nexus 9300-FX3 シリーズスイッチ
10.2(3)F 以降	Cisco Nexus 9300-GX2 シリーズスイッチ
10.4(1)F 以降	Cisco Nexus 9332D-H2R スイッチ
10.4(2)F 以降	Cisco Nexus 93400LD-H1 スイッチ
10.4(3)F 以降	Cisco Nexus 9364C-H1 スイッチ Cisco Nexus 9800 シリーズ スイッチ
10.5(2)F 以降	N9K-X9736C-FX3 ライン カードを搭載した Cisco Nexus 9500 プラットフォーム スイッチ
10.5(3)F 以降	Cisco Nexus 9364E-SG2-Q および 9364E-SG2-O スイッチ

サポートされるリリース	サポートされるプラットフォーム
10.6(1)F 以降	Cisco Nexus 9336C-SE1 スイッチ
10.6(2)F 以降	Cisco N9324C-SE1U、N9348Y2C6D-SE1U、N9396Y12C-SE1、および N9396T12C-SE1 スイッチ

VXLAN NGOAM の構成

Cisco NexusスイッチでVXLAN NGOAM を構成するには、次の手順を実行します。

始める前に

前提条件として、VXLAN の構成が完了していることを確認します。

手順

ステップ 1 NGOAM 機能を有効にするには、グローバル コンフィギュレーション モードで **feature ngoam** コマンドを実行する必要があります。

例：

```
switch# configure terminal
switch(config)# feature ngoam
```

ステップ 2 （任意） **show running-config ngoam** コマンドを実行して、ngoam 設定情報を確認します。

例：

```
switch# show run ngoam
feature ngoam
```

NGOAM プロファイルを構成します

Cisco Nexus スイッチで VXLAN プロファイルを構成するには、次の手順に従います。

始める前に

開始する前に、 **feature ngoam** の構成が完了していることを確認します。

手順

ステップ 1 グローバル コンフィギュレーション モードで **ngoam profile** *profile-id* コマンドを実行し、NGOAM プロファイルを有効にします。

例：

```
switch# configure terminal
switch(config)# feature ngoam
switch(config)# ngoam profile 1
switch(config-ng-oam-profile)#
```

指定できる範囲は 1 ～ 1023 です。Default: NA.

ステップ 2 [**description** | **dot1q** | **flow** | **hop count** | **interface** | **oam-channel 2** | **payload** | **sport**] コマンドを使用して、NGOAM プロファイルを構成するために必要なオプションを設定します。

例：

ngoam プロファイル モードで構成された NGOAM プロファイルのさまざまなオプション。

```
switch(config)# ngoam profile 1
switch(config-ngoam-profile)# oam-channel 2
switch(config-ngoam-profile)#flow forward payload pad 0x2
switch(config-ngoam-profile)#sport 12345, 54321
```

ngoam プロファイル フロー サブモードで構成された NGOAM プロファイルのさまざまなオプション。

```
switch(config)# ngoam profile 1
switch(config-ngoam-profile)# flow forward
switch(config-ng-oam-profile-flow)# oam-channel 2
switch(config-ng-oam-profile-flow)# payload
```

- description: このオプションを使用して、プロファイルの説明を構成します。
- dot1q：dot1q タグを使用してカプセル化を指定するには、このオプションを使用します。
- flow：このオプションを使用して、ngoam フローを構成します。
- hop：このオプションを使用して、ngoam ホップカウントを構成します。範囲：1～255。
- interface：このオプションを使用して、NGOAM 出力インターフェイスを構成します。
- oam-channel：このオプションを使用して、Oam-channel を NVO3 sissa に設定します。
- payload：このオプションを使用して、NGOAM ペイロード構成します。
- sport：このオプションを使用して、NGOAM UDP の送信元ポート範囲を構成します範囲：1 ～ 65535。

ステップ 3 （任意） **show running-config ngoam** コマンドを実行して、ngoam プロファイルの設定情報を確認します。

例：

```
switch# show run ngoam
feature ngoam
ngoam profile 1
oam-channel 2
flow forward payload pad 0x2
```

show running-config コマンド出力の NGOAM 設定の配置が更新されました。以前は、NGOAM 構成はインターフェイス構成の前に表示されていました。Cisco NX-OS リリース以降、10.6(1)F NGOAM 設定は、show running-config コマンド出力のインターフェイス レベルの構成の後に表示されます。

変更前の例：

```
ngoam profile 1
  oam-channel 2
```

```
interface Ethernet1/1
  no switchport
  ip address 60.60.60.1/24
  no shutdown
```

変更後の例：

```
interface Ethernet1/1
  no switchport
  ip address 60.60.60.1/24
  no shutdown

interface loopback10
  vrf member Org1:vrf1
  ipv6 address 2010::10/128
ngoam profile 1
  oam-channel 2
```

障害分離と検証ツールの例

VXLAN ネットワークでは、宛先に到達するためにフレームが通過するスイッチのリストを見つけることが望ましい場合があります。送信元スイッチから宛先スイッチへのループバックテストが失敗した場合、パス内の問題のあるスイッチを見つける必要があります。

障害の切り分けおよび検証ツール（VXLAN OAM ツール）は、次の目的で使用されます。

- IP ネットワークの問題を迅速に特定し、
- VXLAN ネットワーク内のホストと VTEP に到達可能性情報を提供します。

Category	Tools
障害検査	ループバック（ping）メッセージ
障害の隔離	Traceroute および Pathtrace メッセージ

ping メッセージ

ループバック（ping）メッセージは、障害の検証に使用されるユーティリティ ツールです。ループバック メッセージ ユーティリティは、さまざまなエラーやパス障害を検出するために使用されます。

注： **Ping は、 ループバック**と呼ばれることがよくあります。

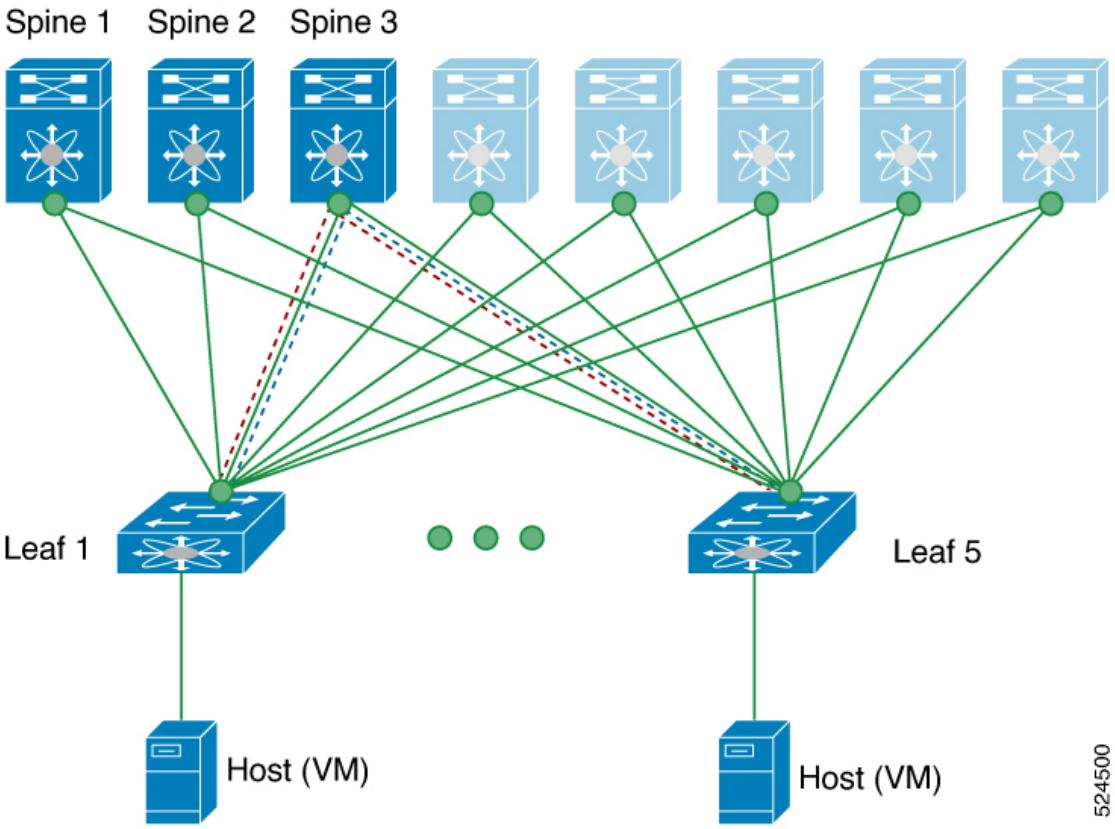
ping 機能のトポロジ

Spine 1、Spine 2、Spine 3 というラベルの付いた 3 つのコア（スパイン）スイッチと 5 つのリーフ スイッチのある Clos トポロジを示します。

ping コマンドを使用する ping メッセージは、これらのネットワーク到達可能性オプションを検証します。

異なる OAM チャネルでリーフ 1（VTEP 1）からリーフ 5（VTEP 2）へのネットワーク到達可能性の検証がどのように行われるかを示します。

ICMP チャネル	NVO3 ドラフト Tissa チャネル
リーフ 1 からリーフ 5 へのループバック メッセージの開始	リーフ 1 からリーフ 5 へのループバック メッセージの開始
ループバック メッセージは、スパイン 3 からの外部ヘッダーに基づいて、VXLAN カプセル化データ パケットとして転送されます。	ループバック メッセージは、スパイン 3 からの外部ヘッダーに基づいて、VXLAN カプセル化データ パケットとして転送されます。
ループバックは、リーフ 5 からのインバンドで処理され、応答されます。	• メッセージが処理され、ペイロードがカプセル化解除されて、ホスト（VM）に送信されます。 • ホストは、受信ペイロードの応答を生成し、リーフ 5 に送信します。 • 応答は、リモート VTEP から受信してインバンドで送信された応答に対して処理されます。
ループバック応答はリーフ 1 に向かうよう処理されます。	ループバック応答はリーフ 1 に向かうよう処理されます。



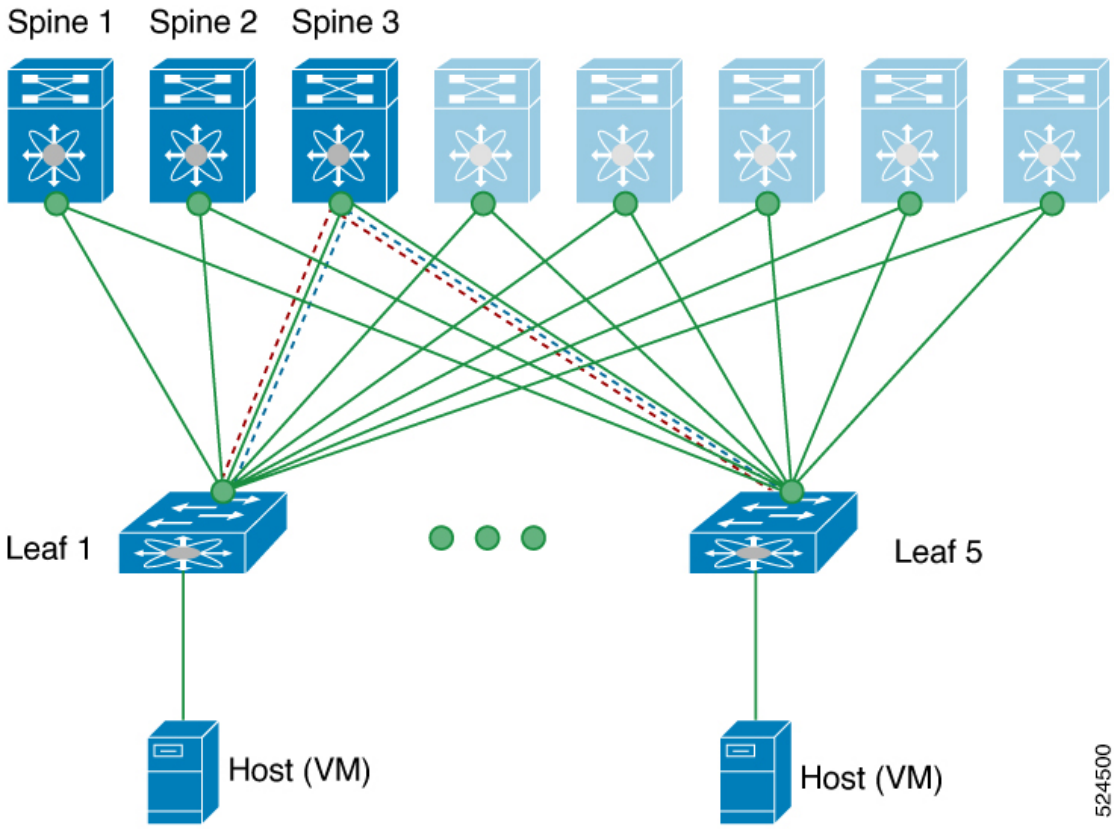
Traceroute メッセージ

Traceroute メッセージは、障害分離に使用されます。このユーティリティは、さまざまなエラーやパス障害をトレースします。

Traceroute の動作は、単一サイトとマルチサイトのシナリオでは異なり、さまざまな出力があります。単一サイトでの Pathtrace 機能 （7 ページ） および マルチサイトでの Traceroute 機能 （8 ページ） セクションで説明されているように、Traceroute のプロセスを理解することが重要です。

単一サイトでの Pathtrace 機能

Spine 1、Spine 2、Spine 3 というラベルの付いた 3 つのコア（スパイン）スイッチと 5 つのリーフ スイッチのある Clos トポロジを示します。



Traceroute メッセージでは、**Traceroute** コマンドを使用して、VXLAN オーバーレイでパケットが通過するパスを検証します。

次に、 VXLAN でカプセル化された ICMP パケット（channel-1）を介したリーフ 1（VTEP 1）からリーフ 5（VTEP 2）への Traceroute メッセージのネットワーク到達可能性の検証を示します。

- traceroute メッセージは、Spine 3 を介して Leaf 1 から Leaf 5 に開始されます。
- Traceroute メッセージは、Spine 3 からの外部ヘッダーに基づいて VXLAN カプセル化データ パケットとして転送されます。
- Traceroute は、Leaf 5 でインバンドで処理および応答されます。
- Traceroute は Leaf 1 で処理されます。

マルチサイトでの Traceroute 機能

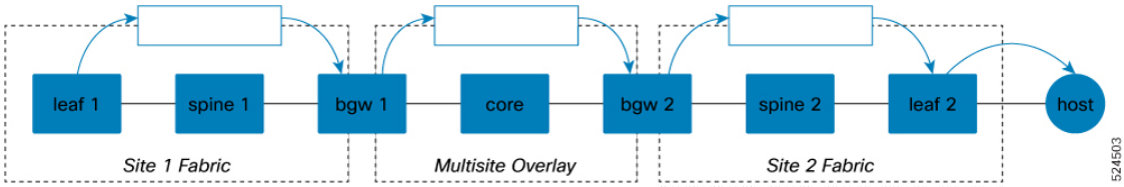
Traceroute： IP

Traceroute（IP）：図が示すように、TTL の有効期限または ACL のヒット数とともに複数のプローブが送信されます。場所は次のとおりです。

- ノードを指す矢印は、トレースがヒットするように見えるホップを示します。

- パイプを指す矢印は、VXLAN でカプセル化されているパケットを表します。

  カプセル化すると、カプセル化によって外部パケットにさらに大きな TTL が追加されるため、パイプからドロップされるまでノードからの応答は表示されません。これは、traceroute が依存している TTL の期限切れは、パイプ内では発生しないことを意味します。



IP Traceroute プロセス

IP traceroute についての説明には、次の手順が含まれます。

- 通常の UDP パケットが開始され、リーフ スイッチの VXLAN 内でカプセル化されます。
- パケットはネットワークを通過し、サイト 1 のボーダー ゲートウェイ（BGW）でカプセル化解除されます。
- サイト 1 の BGW はパケットを受信し、応答を送信します。
- その後、パケットは BGW 1 で再カプセル化され、ネットワークを通過し続けます。
- パケットは BGW 2 でトンネルを出て、別の応答を受信します。
- パケットはもう一度カプセル化され、サイト 2 のリーフから出て、別の応答を求めます。
- 最後に、パケットはリーフに到達し、最後の応答が表示されます。

このシーケンスにより、パケットがさまざまなサイトやネットワーク コンポーネントを通過するときに適切にカプセル化され、またカプセル化解除されることが保証されて、パケットのパスを正確に追跡できます。

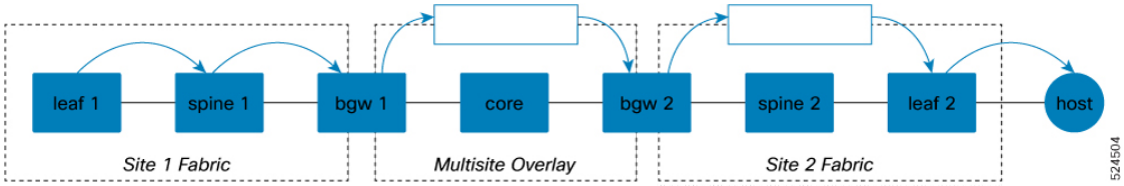
Traceroute - NVE

図に示されているように、NVE Traceroute では、NGOAM は Traceroute が VTEP から生成され、次の手順を実行することを識別します。

- 最初にリモート VTEP までのアンダーレイ ネットワークをトレースします。
- その後、IP traceroute と同様に機能するオーバーレイ Traceroute に切り替わります。
- リモート VTEP の後に、アンダーレイに UDP 要求を使用し、オーバーレイに ICMP 要求を使用します。
- プローブは、リモート VTEP に到達すると VXLAN 内でカプセル化されます。

（注） ローカル ファブリックのプローブは VXLAN にカプセル化されないため、ノードを可視化できます。

- ローカル BGW の後に、プローブがマルチサイトおよびサイト 2 ファブリックのパイプに入ったときの出力は、通常の IP Traceroute に似ています。

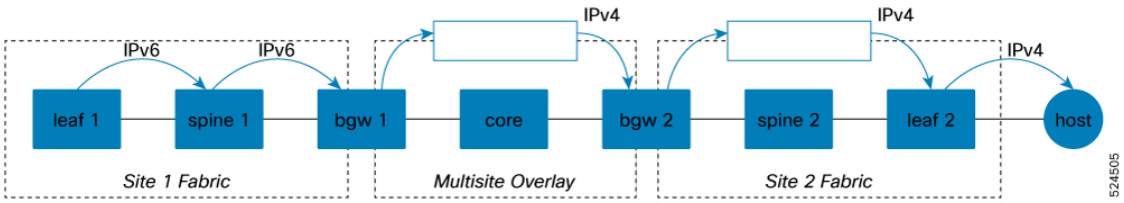


Traceroute（NVE - IPv4 over IPv6）

ハイブリッド トレースでは、アンダーレイ ネットワークとオーバーレイ ネットワーク間の移行のため、IPv6 応答と IPv4 応答が混在します。

図に示されているように、NVE- IPv4 over IPv6 Traceroute では、NGOAM はローカル アンダーレイ ファブリック内に IPv6 としてプローブを生成し、次の手順を実行します。

- ローカル スパインおよび BGW から IPv6 応答を受信します。
- ただし、トレースが BGW に到達すると、NGOAM はオーバーレイ トレースに切り替えます。
- オーバーレイが IPv4 であるため、パケットが効果的に IPv4 になるため、BGW を超えて可視ノードから IPv4 応答を受信します。



Pathtrace メッセージ

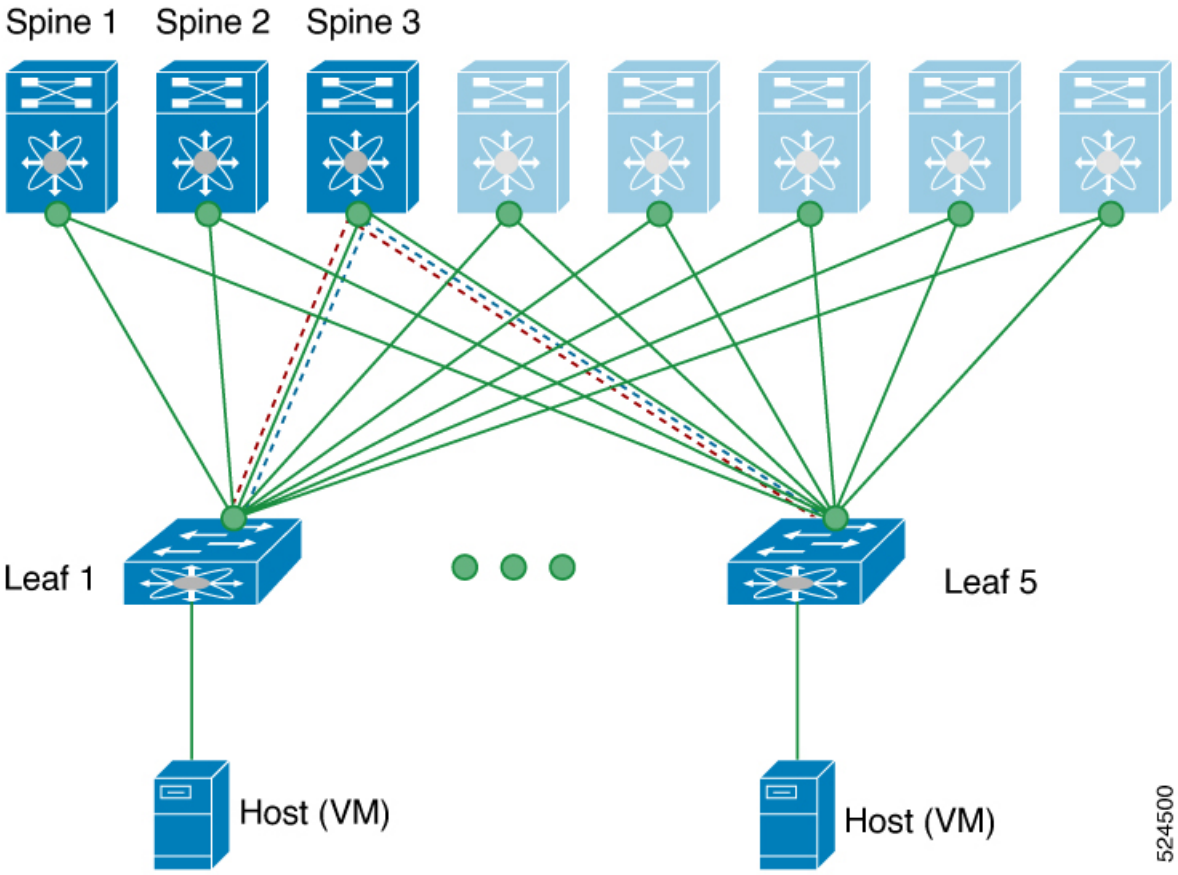
Pathtrace メッセージは、障害分離に使用されます。このユーティリティ ツールは、さまざまなエラーとパス障害をトレースします。

単一サイトでの Pathtrace 機能

Spine 1、Spine 2、Spine 3 というラベルの付いた 3 つのコア（スパイン）スイッチと 5 つのリーフ スイッチのある Clos トポロジを示します。

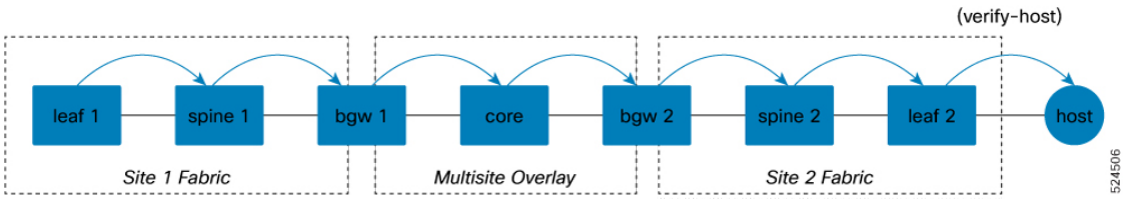
Pathtrace は、VXLAN カプセル化でカプセル化された NVO3 ドラフトの Tissa チャネルを使用してホストに到達し、**Pathtrace** コマンドを使用して VXLAN オーバーレイ内のパケットが通過するパスをトレースします。

NVO3 ドラフトの Tissa チャネルは、入力インターフェイスや出力インターフェイスなど、パスに関する追加情報を提供します。これらのパケットはVTEP で終端し、ホストに到達しません。したがって、VTEP のみが応答します。



マルチサイトでの Pathtrace 機能

図に示されているように、pathtraceはファブリック内の各ノードから応答を生成します。異なるチャネル（NVO3）を使用します。これにより、ファブリック内の VXLAN 対応ノードは、TTL の期限切れではなく ACL のヒットにより特別にパケットを処理できます。これにより、ノードが NGOAM をサポートしている場合は、パケットをキャプチャして処理することが可能になります。また、pathtrace は、BGW 上の NGOAM による特別な処理を受けます。これにより、プローブは次のファブリックに進むように調整されます。



Traceroute および Pathtrace メッセージ

機能	Traceroute メッセージ	Pathtrace メッセージ
使用チャネル	ICMP チャネル	NVO3 ドラフト Tissa チャネル
目的	パケットが接続先に到達するまでにたどるパスの検出による障害分離	インターフェイスの負荷やホップの統計情報など、追加の診断情報を提供
サポートされていないデバイスでの動作	ホップ情報を提供し続ける	単純な traceroute として動作し、ホップ情報のみを提供する

障害の切り分けと検証ツールの注意事項

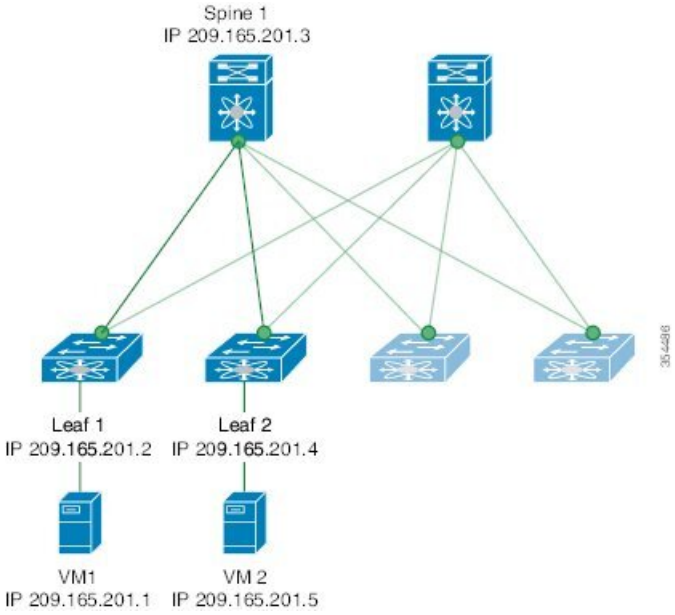
NX-OS リリース9.3(3) 以降、コマンドの **Received** フィールドは、要求がそのノード宛てかどうかに関係なく、**show ngoam pathtrace statistics summary** コマンドが実行されたノードによって受信されたすべてのパストレース要求を示します。

障害分離と検証ツールの例

ping メッセージの例

VXLAN OAM は、スイッチ レベルでホストの可視性を提供し、**ping nve** コマンドを使用してリーフがホストに ping を実行できるようにします。

図 *1 : VXLAN* ネットワーク



次に、チャネル 1（一意のループバック）およびチャネル 2（NVO3 ドラフト Tissa）を使用して、スパイン 1 を介してリーフ 1 から VM2 に ping を実行する例を示します。

- switch# **ping nve ip 209.165.201.5 vrf vni-31000 source 1.1.1.1 verbose**

```
Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'D' - Destination
Unreachable, 'X' - unknown return code, 'm' - malformed request(parameter problem),
'c' - Corrupted Data/Test, '#' - Duplicate response


Sender handle: 34
! sport 40673 size 39,Reply from 209.165.201.5,time = 3 ms
! sport 40673 size 39,Reply from 209.165.201.5,time = 1 ms
! sport 40673 size 39,Reply from 209.165.201.5,time = 1 ms
! sport 40673 size 39,Reply from 209.165.201.5,time = 1 ms
! sport 40673 size 39,Reply from 209.165.201.5,time = 1 ms
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/4/18 ms Total time
elapsed 49 ms
```

（注） 上記の例で使用されている送信元 IP アドレス 1.1.1.1 は、宛先 IP アドレスと同じ VRF のリーフ 1 に構成されているループバックインターフェイスです。たとえば、この例の VRF は vni-31000 です。

- switch# **ping nve ip unknown vrf vni-31000 payload ip 209.165.201.5 209.165.201.4 payload-end verify-host**

```
<snip>
Sender handle: 34
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/4/18 ms Total time
elapsed 49 ms
```

次の例は、NVO3 ドラフト Tissa チャネルを使用して、リーフ 2 からリーフ 1 に MAC ping を実行する方法を示しています。

```
switch# ping nve mac 0050.569a.7418 2901 ethernet 1/51 profile 4 verbose

Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'D' - Destination Unreachable,
 'X' - unknown return code, 'm' - malformed request(parameter problem),
'c' - Corrupted Data/Test, '#' - Duplicate response

Sender handle: 408
!!!!Success rate is 100 percent (5/5), round-trip min/avg/max = 4/4/5 ms Total time
elapsed 104 ms

switch# show run ngoam
feature ngoam ngoam profile 4
oam-channel 2 ngoam install acl
```

Traceroute メッセージの例

次に、スパイン 1 を介してリーフ 1 から VM 2 に traceroute を実行する例を示します。

```
switch# traceroute nve ip 209.165.201.5 vrf vni-31000 source 1.1.1.1 verbose


Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'D' - Destination Unreachable,
 'X' - unknown return code, 'm' - malformed request(parameter problem),
'c' - Corrupted Data/Test, '#' - Duplicate response


Traceroute request to peer ip 209.165.201.4 source ip 209.165.201.2 Sender handle: 36
1 !Reply from 209.165.201.3,time = 1 ms
2 !Reply from 209.165.201.4,time = 2 ms
3 !Reply from 209.165.201.5,time = 1 ms
The following example displays the output of the pathtrace from Leaf 2 to Leaf 1.
switch# pathtrace nve ip 209.165.201.4 vni 31000 verbose
Path trace Request to peer ip 209.165.201.4 source ip 209.165.201.2 Sender handle: 42
Hop Code ReplyIP IngressI/f EgressI/f State
======================================================================
1 !Reply from 209.165.201.3, Eth5/5/1 Eth5/5/2 UP/UP
2 !Reply from 209.165.201.4, Eth1/3 Unknown UP/DOWN
```

Pathtrace メッセージの例

次に、リーフ 2 からリーフ 1 へのペイロードに基づいてパス トレースする例を示します。

```
switch# pathtrace nve ip unknown vrf vni-31000 payload mac-addr 0050.569a.d927
0050.569a.a4fa ip 209.165.201.5 209.165.201.1 port 15334 12769 proto 17 payload-end
Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'D' - Destination Unreachable,
 'X' - unknown return code, 'm' - malformed request(parameter problem),
'c' - Corrupted Data/Test, '#' - Duplicate response

Path trace Request to peer ip 209.165.201.4 source ip 209.165.201.2 Sender handle: 46
Hop Code Reply IngressI/f EgressI/f State
======================================================================
1 !Reply from 209.165.201.3, Eth5/5/1 Eth5/5/2 UP/UP
2 !Reply from 209.165.201.4, Eth1/3 Unknown UP/DOWN
```

（注） 最終宛先までの合計ホップ カウントが 5 を超える場合、パス トレースのデフォルト TTL 値は 5 です。VXLAN OAM パス トレースを完全に終了するには、**max-ttl** オプションを使用します。

次に例を示します。 **pathtracenve ip unknown vrf vrf-vni13001 payload ip 200.1.1.71 200.1.1.23 payload-end verbose max-ttl 10**

次に、NVE MAC に pathtrace する例を示します。

```
switch# pathtrace nve mac 0050.569a.d927 11 payload mac-addr 0050.569a.d927 0050.569a.a4fa
 payload-end vni 31000 verbose

Codes: '!' - success, 'Q' - request not sent, '.' - timeout, 'D' - Destination Unreachable,
 'X' - unknown return code, 'm' - malformed request(parameter problem),
'c' - Corrupted Data/Test, '#' - Duplicate response, 'v' - Other - Use verbose to see
the result

Path trace Request to peer ip 209.165.201.4 source ip 209.165.201.2 Sender handle: 46
Hop Code Reply IngressI/f EgressI/f State
======================================================================
1 !Reply from 209.165.201.3, Eth5/5/1 Eth5/5/2 UP/UP
2 !Reply from 209.165.201.4, Eth1/3 Unknown UP/DOWN
```

（注） 最終宛先までの合計ホップ カウントが 5 を超える場合、パス トレースのデフォルト TTL 値は 5 です。VXLAN OAM path trace を完全に終了するには、**max-ttl** オプションを使用します。

次に例を示します。 **pathtrace nve ip unknown vrf vrf-vni13001 payload ip 200.1.1.71200.1.1.23 payload-end verbose max-ttl 10**

VXLAN EVPN ループの検出と緩和の

ループの原因と影響

ループは通常、ファブリックの南側（アクセス側）の配線が正しくないために、VXLAN EVPN ファブリックで発生します。ブロードキャスト パケットがループを持つネットワークに注入されると、フレームはループ内でブリッジされたままになります。より多くのブロードキャスト フレームがループに入ると、それらが蓄積され、サービスの重大な中断を引き起こす可能性があります。

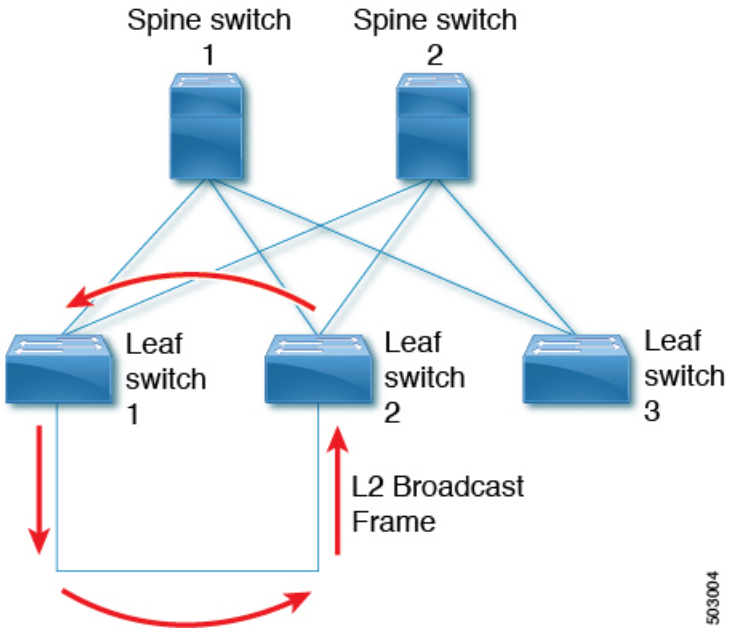
VXLAN EVPN ループの検出と緩和

Cisco NX-OS リリース 9.3(5) では、VXLAN EVPN ループの検出と緩和が導入されています。この機能は、単一の VXLAN EVPN ファブリックまたはマルチサイト環境でレイヤ 2 ループを検出します。ポート/VLAN レベルで動作し、ループが検出された各ポートで VLAN を無効に

します。管理者は、（syslog を介して）条件についても通知されます。このように、この機能により、ネットワークが稼働したままになります。

次の図は、2 つのリーフデバイス（Leaf1 および Leaf2）が南側で直接接続されている EVPN ファブリックを示しています。このトポロジでは、Leaf3 は L2 ブロードキャスト フレームを Leaf1 に転送します。次に、ブロードキャスト フレームは Leaf1 と Leaf2 の間で、南側とファブリックを介して繰り返し転送されます。不正なケーブル接続が修正されるまで、転送が続行されます。

図 *2*: 直接接続された *2* つのリーフ ノード



この機能は、次の 3 つのフェーズで動作します。

1. ループ検出：次の状況でループ検出プローブを送信します。定期的なプローブタスクの一部として、クライアントから要求されたとき、およびポートが起動するとすぐに送信します。

2. ループ緩和：ループが検出されると、ポート上の VLAN をブロックし、次のような syslog メッセージを表示します。

   ```
   2020 Jan 14 09:58:44 Leaf1 %NGOAM-4-SLD_LOOP_DETECTED: Loop detected - Blocking vlan
    1001 :: Eth1/3
   ```

   または

   ```
   2024 Sep  9 15:28:01 Node-11 %ETHPORT-3-IF_ERROR_VLANS_SUSPENDED: VLANs 2704 on
   Interface Ethernet1/49/1 are being suspended. (Reason: SUCCESS)
   ```

   ループは不正なローカル MAC アドレスの学習につながる可能性があるため、このフェーズではローカルおよびリモート MAC アドレスもフラッシュされます。これにより、誤って学習された MAC アドレスが削除されます。

   前の図では、リモート リーフ（Leaf3）の背後にあるホストからのパケットがアクセス側から Leaf1 と Leaf2 の両方に到達できるため、MAC アドレスが誤って学習される可能性が

あります。その結果、ホストは Leaf1 および Leaf2 に対してローカルに誤って表示され、リーフは MAC アドレスを学習します。

3. ループ リカバリ：特定のポートまたは VLAN でループが検出され、リカバリ間隔が経過すると、リカバリ プローブが送信され、ループがまだ存在するかどうかが判断されます。ループから NGAM が回復すると、次のような syslog メッセージが表示されます。

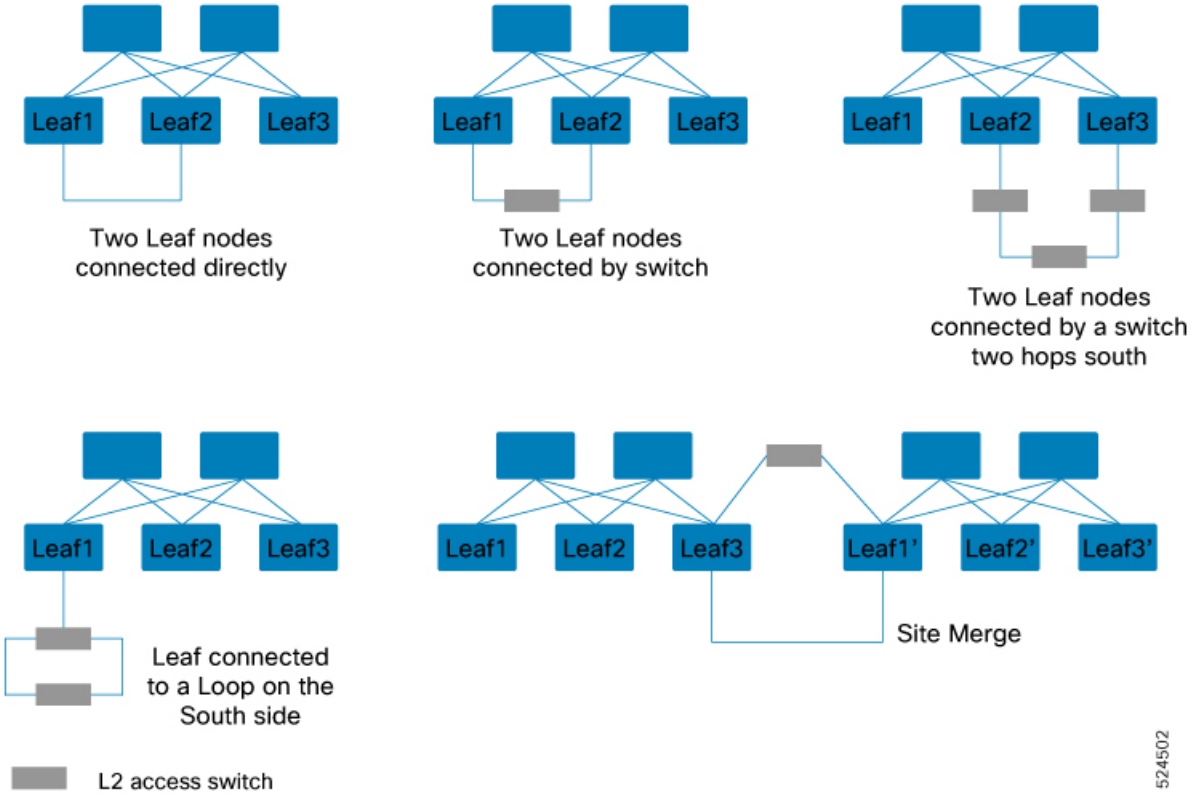
```
2020 Jan 14 09:59:38 Leaf1 %NGOAM-4-SLD_LOOP_GONE: Loop cleared - Enabling vlan 1001

:: Eth1/3
```

または

```
2024 Sep  9 15:24:23 Node-11 %ETHPORT-3-IF_ERROR_VLANS_REMOVED: VLANs 384 on Interface
 Ethernet1/49/1 are removed from suspended state.
```

（注） NGAM のデフォルトのロギングレベルでは、syslog メッセージは生成されません。「logging level ngoam 5」を使用して NGAM のロギングレベルを 5 に変更すると、ループが検出されたときに syslog メッセージが生成されます。

さまざまなループ シナリオ



レイヤ 3 インターフェイス上のサウスバウンド ループ検出

NX-OS リリース 10.4(3)F 以降、Cisco Nexus スイッチは、レイヤ 3（L3）イーサネットおよび L3 ポートチャネルインターフェイス（単一の VXLAN EVPN ファブリックまたはマルチサイト環境にあるもの）でのサウスバウンドループ検出（SLD）をサポートしています。このリリース以前は、SLD 機能はレイヤ 2 インターフェイスでのみサポートされていました。

この機能は、L3 インターフェイスまたはポートチャネルを介して単一のリーフスイッチに接続されているサウスバウンド側（L2 アクセススイッチ）のループを検出します。

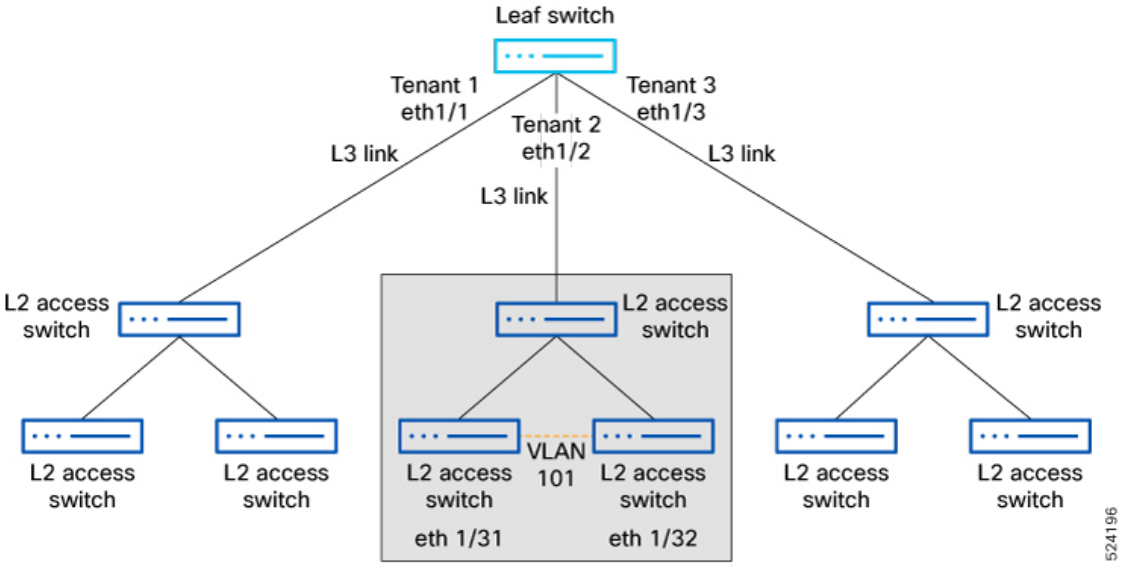
SLD 機能を L3 インターフェイスで有効にすると、この機能は定期的な SLD プローブを送信して、ダウンストリームテナントのレイヤ 2 ドメイン内のループを検出します。ユーザーがダウンストリーム L2 ドメインの状態を修正するためのアクションを実行するまで、ループのモニターを継続し、検出時には L3インターフェイスをブロックします。

レイヤ 3 インターフェイスでの SLD の機能

- 単一の L3 アタッチテナントを分離して、コントロールプレーンポリシングの輻輳が原因でストームの影響が単一の L3境界を超えて伝播するのを防ぎます。
- 発信元 NGOAM プローブの受信によってループが検出された場合、ダウンストリームの L2 ループを検出し、アタッチされた L3 インターフェイスまたは L3 ポートチャネルをブロックします。
- 発信元 NGOAM プローブがループを検出しなくなれば、L3 ポートのブロックを解除します。

レイヤ 3 インターフェイス上の SLD のトポロジの概要

次の図は、3 つの VRF（テナント 1、テナント 2、およびテナント 3）で構成されたリーフスイッチを備えている EVPN ファブリックを示しています。これらの VRF は、異なる L3 ポートとそれぞれの L3 インターフェイスを使用して、サウス側の L2アクセススイッチに接続されます。



この機能は、次の 3 つのフェーズで動作します。

- **ループ検出**：SLD L3 機能は、定期的にプローブを送信して、ダウンストリームテナントのレイヤ 2 ドメイン（L2 アクセススイッチ）のループを検出します。

  SLD は、次の状況でループ検出プローブを送信します。クライアントから要求されたとき、定期的なプローブタスクの一部として、および何らかのポートが起動したときです。

  例：ローカル VLAN 101 で STP を無効にしているときに、ケーブル接続の間違いにより、テナント 2 で誤ってブリッジループが形成されたとします。これにより、Eth1/2 への ARP ストームがトリガーされ、CoPP クラスのノーマルポリサー全体が消費され、テナント 1 とテナント 3 で CoPP ポリサーが飽和状態になります。

  ```
  2024 Jun 27 02:34:39 tenant2 %L2FM-2-L2FM_CONTINUOUS_MAC_MOVE: Mac
  Address (f80f.6f96.a127) in Vlan 101 is moving continuously. Mac
  moved between Eth1/32 to Eth1/31. Please enable 'logging level l2fm
  4' for verbose output.
  ```

- **ループの緩和**：ループが検出されたときに L3 ポートをブロックし、ループ検出とポートステータスの変更を示す次のような syslogメッセージを表示します。

  ```
  2024 Jun 27 02:37:50 leaf %ETHPORT-5-IF_DOWN_NONE: Interface Ethernet1/2 is down
  (None)
  2024 Jun 27 02:37:50 leaf %ETHPORT-5-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/2
  is down (Error disabled. Reason:error)
  2024 Jun 27 02:38:52 leaf %ETHPORT-5-IF_ERRDIS_RECOVERY: Interface Ethernet1/2 is
  being recovered from error disabled state (Last Reason:error)
  2024 Jun 27 02:38:54 leaf %ETHPORT-5-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/2
  is down (Error disabled. Reason:error)
  !
  leaf# show ngoam loop-detection status l3
  Port       Status       NumLoops      DetectionTime             ClearedTime
  ==========================================================================================

  Eth1/2     BLOCKED      2             Tue Jun 27 02:38:54 2024  Tue Jun 27 02:38:52
   2024
  ```

各プローブのエラーリカバリ間隔の経過後、ブロックされていた L3 ポートがアップ状態になり、プローブを送信し、ループを再確認します。これで、Eth1/2 L3インターフェイスが**ブロック中**状態から**転送中**状態に移行します。プローブはループをチェックし、ループがまだ存在する場合は、eth1/2 L3 インターフェイスを **ブロック** 状態に戻します。このプロセスは、ユーザーが L2ドメイン内のブリッジングループを修正するまで続きます。

次の出力例は、生成されたプローブに基づいて、状態（ブロッキングおよびブロック解除）を示しています。

```
2024 Jun 27 20:26:56 leaf %NGOAM-4-SLD_L3_LOOP_DETECTED: Loop detected - Blocking
port Eth1/2
2024 Jun 27 20:26:56 leaf %ETHPORT-5-IF_DOWN_NONE: Interface Ethernet1/2 is down
(None)
2024 Jun 27 20:26:56 leaf %ETHPORT-5-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/2
is down (Error disabled. Reason:error)
2024 Jun 27 20:27:58 leaf %ETHPORT-5-IF_ERRDIS_RECOVERY: Interface Ethernet1/2 is
being recovered from error disabled state (Last Reason:error)
2024 Jun 27 20:27:58 leaf %NGOAM-4-SLD_L3_LOOP_GONE: Loop cleared - Enabling port
Eth1/2
2024 Jun 27 20:28:00 leaf %NGOAM-4-SLD_L3_LOOP_DETECTED: Loop detected - Blocking
port Eth1/2
2024 Jun 27 20:28:01 leaf %ETHPORT-5-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/2
is down (Error disabled. Reason:error)
```

- **ループリカバリ**：ケーブルエラーを直すと、サウスバウンド側のループは解消されます。リカバリ間隔が経過すると、リーフスイッチの L3 インターフェイスからリカバリプローブが送信され、ループがまだ存在するかどうかが判断されます。ループが解決されていれば、ポートは転送中状態のままになり、次の syslog メッセージが生成されます。

```
2024 Jun 27 22:39:26 tenant2 %ETHPORT-5-IF_DOWN_ADMIN_DOWN: Interface Ethernet1/32
is down (Administratively down)
2024 Jun 27 22:39:56 tenant2 %ETHPORT-5-SPEED: Interface Ethernet1/2, operational
speed changed to 10 Gbps
2024 Jun 27 22:39:56 tenant2 %ETHPORT-5-IF_DUPLEX: Interface Ethernet1/2, operational
 duplex mode changed to Full
2024 Jun 27 22:39:56 tenant2 %ETHPORT-5-IF_RX_FLOW_CONTROL: Interface Ethernet1/2,
operational Receive Flow Control state changed to off
2024 Jun 27 22:39:56 tenant2 %ETHPORT-5-IF_TX_FLOW_CONTROL: Interface Ethernet1/2,
operational Transmit Flow Control state changed to off
2024 Jun 27 22:39:56 tenant2 %ETHPORT-5-IF_UP: Interface Ethernet1/17 is up
2024 Jun 27 22:41:03 tenant2 %VSHD-5-VSHD_SYSLOG_CONFIG_I: Configured from vty by
admin on 10.82.195.201@pts/2
```

（注） NGAM のデフォルトのロギングレベルでは、syslog メッセージは生成されません。「logging level ngoam 5」を使用して NGAM のロギングレベルを 5 に変更すると、ループが検出されたときに syslog メッセージが生成されます。

L2 および L3 SLD 機能の比較

機能	L2 インターフェイスの SLD	L3 インターフェイスの SLD
運用レベル	ポートおよび VLAN レベル	イーサネットおよびL3 ポートチャネル

機能	L2 インターフェイスの SLD	L3 インターフェイスの SLD
環境	シングルサイトとマルチサイト	シングルサイトとマルチサイト
ループ検出	特定のポートまたは VLAN のループを検出します。	ダウンストリーム L2 ループを検出し、L3 インターフェイスまたは L3 ポートチャネルをブロックします。
ループの緩和	ループが検出されると、ポート上の VLAN をブロックし、syslog メッセージを表示する	単一の L3 アタッチテナントを分離することにより、共有 CoPP ポリサーリソースが消費されてストームの影響が単一の L3 境界を超えて伝播しないようにする
ループ ブロック	サウスバウンド ループを解消する	ストーム関連トラフィックを遮断することで、検出されたループがコントロール プレーンに影響を与えないように分離
ループ後のリカバリ	ループがクリアされたら、リカバリプローブを送信し、VLAN を再度有効にし、syslog メッセージをログに記録	NGOAM プロセスが NGOAM プローブを認識しなくなった場合は、リカバリ プローブを送信し、ポートまたはイーサネットインターフェイスを再度有効にし、ループがクリアされていたら syslog メッセージをログに記録

VXLAN EVPN ループの検出と緩和のガイドラインと制限事項

VXLAN EVPN ループの検出と緩和には、次のガイドラインと制限事項があります。

- VXLAN EVPN ループの検出と緩和は、STP および STP なしの両方の環境でサポートされます。
- VXLAN EVPN マルチサイト展開のサイト間でループを検出できるようにするには、この機能が展開されているサイト内のすべての境界ゲートウェイで **ngoam loop-detection** コマンドを設定する必要があります。
- VXLAN EVPN ループの検出と緩和は、次の機能ではサポートされません。
  - プライベート VLAN
  - VLAN 変換
  - ESI ベースのマルチホーミング

- VXLAN クロス コネクト
- Q-in-VNI
- EVPN セグメント ルーティング（レイヤ2）

（注） これらの機能が設定されたポートまたはVLANは、VXLAN EVPN ループの検出および緩和から除外する必要があります。これらを除外するには、**disable**{**vlan** *vlan-range*} [**port** *port-range*] コマンドを使用できます。

VXLAN EVPN ループ検出と緩和がサポートされるプラットフォームとリリース

サポートされるリリース	サポートされるプラットフォーム
9.3(5) 以降	Cisco Nexus 9300-FX/FX2 シリーズ スイッチ X97160YC-EX および 9700-FX ライン カード搭載の Cisco Nexus 9500 プラットフォーム スイッチ
10.1(1) 以降	Cisco Nexus 9300-FX3/GX シリーズスイッチ
10.2(3)F 以降	Cisco Nexus 9300-GX2 シリーズスイッチ
10.4(1)F 以降	Cisco Nexus 9332D-H2R シリーズスイッチ
10.4(2)F 以降	Cisco Nexus 93400LD-H1 シリーズスイッチ
10.4(3)F 以降	Cisco Nexus 9364C-H1 シリーズスイッチ
10.5(2)F 以降	9700-FX3 ライン カード搭載の Cisco Nexus 9500 プラットフォーム スイッチ

L3 インターフェイス上の SLD のガイドラインおよび制限事項

- SLDは、L3イーサネットおよびL3ポートチャネルインターフェイスでのみサポートされます。L3 サブインターフェイスではサポートされていません。

L3 インターフェイスの SLD でサポートされるプラットフォームとリリース

リリース	プラットフォーム
10.4(3)F 以降	Cisco Nexus 9300-FX/FX2/GX/GX2/H2R/H1 シリーズ スイッチ 9700-FX/GX ライン カード搭載の Cisco Nexus 9500 プラットフォーム スイッチ
10.5(2)F 以降	9700-FX3 ライン カード搭載の Cisco Nexus 9500 プラットフォーム スイッチ

NGOAM サウスバウンド ループ検出の前提条件

作業を開始する前に、次を確認してください。

- NGOAM 機能を有効にします。
- TCAM ing-sup リージョン用のスペースを作成するには、次の **hardware access-list tcam region ing-sup 768** コマンドを使用します。

（注）
- ing-sup リージョンの割り当てを増やす前に、追加の TCAM エントリが解放されていることを確認します。
- TCAM リージョンを設定するには、ノードをリブートする必要があります。

レイヤ 2 インターフェイス上の NGOAM サウスバウンドループ検出の構成

NGOAM サウスバウンドループの検出と緩和を設定するには、次の手順に従います。

手順

ステップ 1 Run the [**no**] **ngoam loop-detection** command in global configuration mode, to enable NGOAM Southbound loop detection and mitigation for all VLANs or ports.

例：

```
switch# configure terminal
switch(config)# ngoam loop-detection
switch(config-ng-oam-loop-detection)#
```

この機能はデフォルトで無効に設定されています。

このコマンドの **no** 形式は、NGOAM サウスバウンドループ検出と緩和を無効にします。

ステップ 2 （任意） [**no**] **disable** {**vlan** *vlan-range*} [**port** *port-range*]コマンドを実行して、特定の VLAN またはポートの NGOAM サウスバウンドループの検出および緩和を無効にし、ループ検出されたポートを起動します。

例：

特定の VLAN ポートでディセーブルにします：

```
switch(config-ng-oam-loop-detection)# disable vlan 1200 port ethernet 1/1
```

特定の VLAN での無効化：

```
switch(config-ng-oam-loop-detection)# disable vlan 1300
```

このコマンドの **no** 形式は、これらの VLAN またはポートのアクティブ モニタリングを再開します。

ステップ 3 （任意） [**no**] **periodic-probe-interval** *value* コマンドを実行して、定期的なループ検出プローブの送信頻度を指定します。

例：

```
switch(config-ng-oam-loop-detection)# periodic-probe-interval 200
```

範囲： 60〜3600 秒（60 分）。デフォルト: 300秒(5分)。

ステップ 4 （任意） [**no**] **port-recovery-interval** *value* コマンドを実行して、ポートまたは VLAN がシャットダウンされたときにリカバリ プローブが送信される頻度を指定します。

例：

```
switch(config-ng-oam-loop-detection)# port-recovery-interval 300
```

範囲： 300〜3600 秒（60 分）。デフォルト値： 600 秒（10 分）。

ステップ 5 （任意） **show ngoam loop-detection summary** コマンドを実行してループ検出の構成と現在のループの概要を確認します。

例：

```
switch# show ngoam loop-detection summary
Loop detection:enabled
Periodic probe interval: 200
Port recovery interval: 300
Number of vlans: 1
Number of ports: 1
Number of loops: 1
Number of ports blocked: 1
Number of vlans disabled: 0
Number of ports disabled: 0
Total number of probes sent: 214
Total number of probes received: 102
Next probe window start: Thu May 14 15:14:23 2020 (0 seconds)
Next recovery window start: Thu May 14 15:54:23 2020 (126 seconds)
```

次のタスク

スパインの QoS ポリシーを設定します。（構成例については、NGOAM サウスバウンドループの検出と緩和の構成例 （27 ページ）を参照して ください）。

レイヤ 3 インターフェイス上の NGOAM サウスバウンドループ検出の構成

イーサネットおよびL3 ポートチャネルインターフェイスでNGOAM サウスバウンドループ検出を有効にするには、次の手順を実行します。

手順

ステップ 1 Run the [**no**] **ngoam loop-detection** command in global configuration mode, to enable NGOAM Southbound loop detection and mitigation for all VLANs or ports.

例：

```
switch# config terminal
switch(config)# ngoam loop-detection
switch(config-ng-oam-loop-detection)#
```

この機能はデフォルトで無効に設定されています。

ステップ 2 [**no**] **l3 ethernet port** *port-range*コマンドを実行してイーサネット インターフェイスで L3 ループ検出を有効にします。

例：

```
switch(config-ng-oam-loop-detection)# l3 ethernet port Eth1/49
```

イーサネット インターフェイスで L3 ループ検出を無効にするには、このコマンドの **no** 形式を使用します。

ステップ 3 [**no**] **l3 port-channel port** *port-range*コマンドを実行してポートチャネル インターフェイスで L3 ループ検出を有効にします。

例：

```
switch(config-ng-oam-loop-detection)# l3 port-channel port port-channel1
```

ポートチャネル インターフェイスでL3 ループ検出を無効にするには、このコマンドの **no** 形式を使用します。

ステップ 4 （任意） **show ngoam loop-detection status l3** コマンドを実行して、L3 インターフェイスで検出されたループを確認します。

例：

```
switch# show ngoam loop-detection status l3
Port       Status       NumLoops      DetectionTime             ClearedTime
================================================================================================
 Eth1/2     BLOCKED      2             Tue Jun 25 02:38:54 2024  Tue Jun 25 02:38:52 2024
```

ステップ 5 （任意） **show run ngoam** コマンドを実行してループ検出の構成と現在のループの概要を確認します。

例：

```
switch# show run ngoam
ngoam loop-detection
  periodic-probe-interval 60
```

```
  port-recovery-interval 600
  l3 ethernet port Ethernet1/1-3
!
2024 Jun 25 02:37:50 switch %ETHPORT-5-IF_DOWN_NONE: Interface Ethernet1/2 is down (None)
2024 Jun 25 02:37:50 switch %ETHPORT-5-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/2 is down (Error
 disabled. Reason:error)
2024 Jun 25 02:38:52 switch %ETHPORT-5-IF_ERRDIS_RECOVERY: Interface Ethernet1/2 is being recovered
 from error disabled state (Last Reason:error)
2024 Jun 25 02:38:54 switch %ETHPORT-5-IF_DOWN_ERROR_DISABLED: Interface Ethernet1/2 is down (Error
 disabled. Reason:error)
```

ループの検出とオンデマンドでのポートの呼び出し

ループを検出するか、ブロックされたポートをオンデマンドで起動するには、この項の手順に従います。

手順

ステップ 1 （任意） コマンドを実行し **ngoam loop-detection probevlan** て**port** 、 指定された VLAN またはポートでループ検出プローブを送信します。

例：

```
switch# ngoam loop-detection probe vlan 1200 port ethernet 1/1
```

このコマンドは、プローブが正常に送信されたかどうかを確認するための通知も送信します。

ステップ 2 （任意） 以前にブロックされたVLANまたはポートを起動するには、こちら **ngoam loop-detection bringup** [**vlan** *vlan-range*] [**port** *port-range*] コマンドを実行します。

例：

```
switch# ngoam loop-detection bringup vlan 1200 port ethernet 1/1
```

また、このコマンドを実行すると、NGOAM にスタックしているエントリがクリアされます。

（注）
ループが解消されてからポートが起動するまでに、最大で 2 つのポート回復インターバルが必要です。**ngoam loop-detection bringup vlan** {**vlan** *vlan-range*} [**port** *port-range*] コマンドを使用して手動でタイマーを上書きすることで、リカバリを高速化できます。

ステップ 3 （任意） **show ngoam loop-detection status** [**history**] [**vlan** *vlan-range*] [**port** *port-range*] コマンドを実行し、**history** オプションを指定した場合と指定しない場合の、VLAN またはポートのループ検出ステータスを確認します。

例：

履歴 オプションなし

```
switch# show ngoam loop-detection status
VlanId Port   Status     NumLoops  Detection Time                 ClearedTime
====== ====== ========== ========= ============================== ===============
```

```
100    Eth1/3 BLOCKED     1        Tue Apr 14 20:07:50.313 2020  Never
```

履歴 オプションあり

```
switch# show ngoam loop-detection status history
VlanId Port   Status     NumLoops  Detection Time                ClearedTime
====== ====== ========== ========= ============================= ===============
100    Eth1/3 BLOCKED     1        Tue Apr 14 20:07:50.313 2020  Never
200    Eth1/2 FORWARDING  1        Tue Apr 14 21:19:52.215 2020  May 11 21:30:54.830 2020
```

ステータスは、次のいずれかになります。

- **BLOCKED**：ループが検出されたため、VLAN またはポートがシャットダウンされました。
- **FORWARDING**：ループが検出されず、VLAN またはポートが動作しています。
- **RECOVERING**：以前に検出されたループがまだ存在するかどうかを判断するために、回復プローブが送信されています。

history オプションは、ブロックされたポート、転送中のポート、および回復中のポートを表示します。**history** オプションを指定しない場合、コマンドはブロックされたポートと回復中のポートのみを表示します。

NGOAM サウスバウンドループの検出と緩和の構成例

次に、スパインに QoS ポリシーを設定し、ループ検出が有効なリーフが接続されているすべてのスパイン インターフェイスに適用する例を示します。

```
class-map type qos match-any Spine-DSCP56
match dscp 56
policy-map type qos Spine-DSCP56
class Spine-DSCP56
set qos-group 7

interface Ethernet1/31
mtu 9216
no link dfe adaptive-tuning
service-policy type qos input Spine-DSCP5663
no ip redirects
ip address 27.4.1.2/24
ip router ospf 200 area 0.0.0.0
ip pim sparse-mode
no shutdown
```

翻訳について
このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。