VXLAN EVPN トラフィック エンジニアリング：マルチサイト出力ロードバランシングの構成

この章では、Cisco NX-OS デバイスでVXLAN EVPN トラフィック エンジニアリング（TE）：マルチサイト出力ロードバランシング機能を構成する方法を説明します。

この章は、次の項で構成されています。

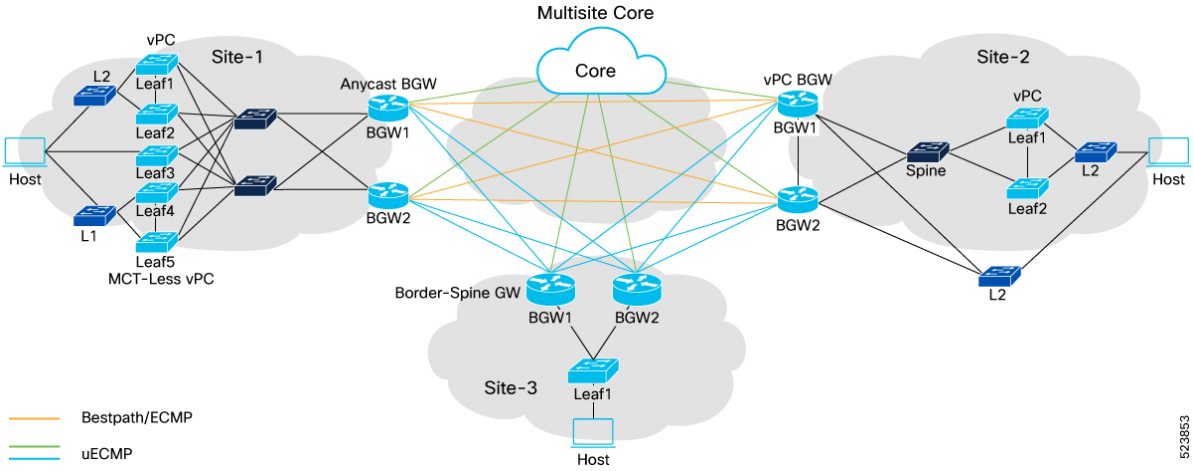
- VXLAN EVPN TE - マルチサイト出力ロードバランシングの概要 （1 ページ）
- VXLAN EVPN TE - マルチサイト出力ロードバランシングの注意事項および制限事項 （2 ページ）
- VXLAN EVPN TE：マルチサイト出力ロードバランシングの構成 （4 ページ）
- VXLAN EVPN TE の確認：マルチサイト出力ロード バランシング構成の確認 （16 ページ）
- VXLAN EVPN TE の構成例：マルチサイト出力ロードバランシング （18 ページ）

VXLAN EVPN TE - マルチサイト出力ロードバランシングの概要

VXLAN EVPN TE - マルチサイト出力ロードバランシング機能は、トラフィック ステアリングを促進し、マルチサイト リンクを介して異なるファブリック間で送信されるデータのロードバランシングを可能にします。

トラフィック エンジニアリングおよびロードバランシング機能は、通常サイト間ネットワーク（ISN）と呼ばれる IP ベースのアンダーレイ ネットワーク全体で動作します。したがって、これは基本的に、アンダーレイを介して送信される任意のオーバーレイ カプセル化トラフィックに適用できる IP トラフィック エンジニアリングとして提供されます。

VXLAN EVPN TE - マルチサイト出力ロードバランシングは、通常、データセンター（DC）間リンクの使用率を向上させます。



上記のトポロジは、同じ Multi-Site ドメインの 3 つの VXLAN EVPN ファブリック部分を示しています。各ファブリックは、基本的にネットワーク インフラストラクチャの残りの部分とのファブリックのインターフェイスを表すボーダー ゲートウェイ デバイス（BGW）のローカル階層を介してリモートサイトに接続します。エニーキャスト BGW、vPC BGW、ボーダー ゲートウェイ スピンなど、さまざまなタイプの BGW は、この展開内で共存できます。これらは、BGW 間の直接リンクや汎用コア インフラストラクチャ（ISN）など、さまざまな接続オプションを介して相互接続できます。

- 通常、オレンジ色で示されたパスは、サイト 1 とサイト 2 に属するエンドポイント間のサイト間通信のベスト パスまたはマルチパスとして使用されます。
- VXLAN EVPN TE - マルチサイト出力ロードバランシング機能を有効にすることで、トラフィック分散用の追加パスをベストパス以外に使用できます。これには、サイト 3 と呼ばれる中間ロケーションを経由するルートなどの代替ルート（青色で強調表示）や、一般的なコア インフラストラクチャを通過するパス（緑色で強調表示）が含まれます。これらの代替パスは、不等コスト マルチパス（uECMP）または重み付け不等コスト マルチパス（wuECMP）の一部として使用できます。

VXLAN EVPN TE - マルチサイト出力ロードバランシングの注意事項および制限事項

- Cisco NX-OS リリース 10.4(3)F 以降、VXLAN EVPN TE - マルチサイト出力ロードバランシング機能が Cisco Nexus 9300 FX/FX2/FX3/GX/GX2 スイッチ、および 9700-FX/GX ラインカードでサポートされます。ただし、現在サポートされているのは BGP ベースのアンダーレイ ルーティングのみです。
- VXLAN EVPN TE：マルチサイト出力ロードバランシング機能は、9500-FM-E ファブリックモジュールを搭載した Cisco Nexus 9500 モジュラ スイッチではサポートされていません。

- VXLAN EVPN TE：マルチサイト出力ロードバランシングは、次の機能をサポートします。
  - リモート サイトへのすべてがベスト パスではない可能性があるアンダーレイ パス間での出力トラフィックのロード バランシング（LB）。
  - 特定の「重み」（または負荷）を各マルチサイトパスに関連付けるための明示的および自動 LB ポリシー。これには、重み付け等コスト マルチパス（wECMP）と 重み付け不等コスト マルチパス（wuECMP）の 2 つのオプションが付属しています。
  - BGP ベース アンダーレイ ルーティング構
  - AS-Path ベースの uECMP パス選択。
  - アンダーレイのレイヤ 3 ユニキャスト uECMP/wuECMP。
  - レイヤ 3（EVPN タイプ 5）およびレイヤ 2（EVPN タイプ 2）オーバーレイ ユニキャスト ECMP/wuECMP。
  - BUM 転送：
    - BUM トラフィックは、出力ロードバランシングの対象になりません。
    - ただし、入力複製がサイト間での BUM 転送（DCI IR）に使用される場合、アンダーレイのネクストホップ パスの選択は、マルチパス セットの出力ロードバランシング パスの一部から 1 つを使用できるため、アンダーレイの uECMP の利点があります。
  - ソフトウェア転送：ソフトウェア転送のベースライン ケースでは、レイヤ2とレイヤ 3 の両方のトラフィックが EVPN VXLAN マルチサイト設定で uECMP/wuECMP をサポートします。
  - CloudSec（PIP ネクストホップ）。
  - ファイアウォール クラスタリング（PIP ネクストホップ）。
  - VNI をダウンストリームします。ただし、 **dci-advertise-pip** を使用した DSVNI はサポートされていません。

（注） ネクストホップとして（マルチサイト VIP ではなく）PIP を使用してタイプ 2 およびタイプ 5 EVPN プレフィックスのアドバタイズを開始するには、BGW で **dci-advertise-pip** コマンドが必要です。オーバーレイとアンダーレイの動的重み（wuECMP）、およびアンダーレイの AIGP （34 ページ） セクションでは、不等コスト ネクストホップ PIP アドレス宛てのオーバーレイ トラフィックに対して wuECMP ロードバランシングを実行する必要がある理由について詳しく説明します。

  - VXLAN OAM。

- ポリシーベース ルーティング（PBR）

- VXLAN EVPN TE - マルチサイト出力ロードバランシングは、次の機能をサポートしていません。
  - IPv6 アンダーレイ を使用した VXLAN。
  - 重み付き ECMP/uECMP を使用した **hardware profile ecmp resilient** 構成。
  - マルチキャスト オーバーレイ トラフィック。
  - wuECMP は、9700-FX ライン カードを搭載した Cisco Nexus 9500 シリーズ スイッチではサポートされません。

- Cisco NX-OS リリース 10.4(3)F から以前のリリースにダウングレードする場合、VXLAN EVPN TE - マルチサイト出力ロードバランシング構成が削除されていることを確認します。

VXLAN EVPN TE：マルチサイト出力ロードバランシングの構成

次に、マルチサイト出力ロードバランシングを有効にするための 3 つの主な構成手順を示します。

（注） 次の構成を BGW にのみ適用します。

1. **フィルタ ポリシーの作成**：これは、同じマルチパス セットのアンダーレイ パス部分でトラフィックを分散するリモート宛先オーバーレイ ネクストホップ IP アドレスを識別するために必要です。このアドレスは、リモート BGW の共通のマルチサイト VIP または一意の PIP である可能性があります（**dci-advertise-pip** が構成されている場合）。

2. **マルチパス ポリシーの作成**：このポリシーは、アンダーレイ パスをマルチパス セットの一部として分類する基準を定義するように構成されます。この定義により、静的または動的な重みを持つ uECMP や wuECMP など、複数のユースケースのプロビジョニングが可能になります。具体的には、複数のオーバーレイ ネクストホップが存在する場合（BGW の PIP アドレスなど）、ネクストホップ アドレスの選択を含めるように wuECMP を拡張して、「マルチレベル」の負荷分散効果を作成できます。

3. **ELB VRF での解決の有効化**：デフォルト VRF のルーティング テーブルではなく、**egress-loadbalance-resolution- VRF** のアンダーレイ テーブルを使用して、宛先オーバーレイのネクストホップ IP アドレスに到達するためのアンダーレイ パスの解決を有効にします。

egress-loadbalance-resolution- VRF は、自動的に作成される新しい内部 VRF です。この VRF は構成できず、削除できません。

VXLAN EVPN TE - マルチサイト出力ロードバランシング機能が構成されている場合：

- アンダーレイ プロトコル（現在は BGP のみ）ルートが追加でインポートされ、このテーブルにインストールされます。
- オーバーレイのネクストホップ解決は、デフォルトのテーブルではなく、このテーブルを介して実行されます。
- これにより、デフォルト VRF にインストールされたベスト パスに加えて、サイト間通信により多くのアンダーレイ パスを使用できます。

アンダーレイの出力ロード バランス フィルタ ポリシーの作成

対象のリモート VTEP など、複数のアンダーレイ パス間での出力ロードバランシングを必要とするアンダーレイ ルートを識別するようにフィルタ ポリシーを構成できます。

BGW で出力ロードバランシング フィルタ ポリシーを構成するには、次の手順を実行します。

手順の概要

1. **configure terminal**
2. **ip prefix-list** *prefix-list-name* **seq** *value* **permit** *nexthop-ipaddress*
3. **route-map** *route-map-name*
4. アンダーレイ ルートは、コミュニティ属性を使用するか、以下に示すようにプレフィックス リストを介して照合されます。
   - **match ip address prefix-list** *prefix-list-name*

     または

   - **match community** *community-list*
5. **exit**
6. **router bgp** *as-number*
7. **address-family ipv4 unicast**
8. **[no] load-balance egress filter-policy route-map** *route-map-name*

手順の詳細

手順

	コマンドまたはアクション	目的

ステップ 1	**configure terminal**
	例：
	`switch# config terminal`
	`switch(config)#`
		グローバル コンフィギュレーション モードを開始します。

	コマンドまたはアクション	目的
ステップ 2	**ip prefix-list** *prefix-list-name* **seq** *value* **permit** *nexthop-ipaddress* 例： `switch(config)# ip prefix-list remote_nexthop seq 5 permit 10.10.112.1/32`	リモート ネクストホップと一致するようにプレフィックス リストを構成します。

ステップ 3	**route-map** *route-map-name*
	例：
	`switch(config)# route-map ROUTE_MAP_1`
	`switch(config-route-map)#`
		出力ロード バランス ロジックは、ルート マップを使用してフィルタ ポリシーで言及されている対象のネクストホップまたは VTEP ルートにのみ適用されます。

ステップ 4	アンダーレイ ルートは、コミュニティ属性を使用するか、以下に示すようにプレフィックス リストを介して照合されます。
	• **match ip address prefix-list** *prefix-list-name*
	または
	• **match community** *community-list*
	例：
	`switch(config-route-map)# match ip address prefix-list remote_nexthop`
	または
	`switch(config-route-map)# match community BGPCommunity`
		アンダーレイ ルートは、プレフィックスリストまたはコミュニティ リストを使用して照合されます。

ステップ 5	**exit**
	例：
	`switch(config-route-map)# exit`
	`switch(config)#`
		BGP ルータ コンフィギュレーション モードを開始します。

ステップ 6	**router bgp** *as-number*
	例：
	`switch(config)# router bgp 65001`
	`switch(config-router)#`
		BGP ルータ コンフィギュレーション モードを開始します。

ステップ 7	**address-family ipv4 unicast**
	例：
	`switch(config-router)# address-family ipv4 unicast`
	`switch(config-router-af)#`
		IPv4ユニキャストアドレスファミリを構成します。

ステップ 8	**[no] load-balance egress filter-policy route-map** *route-map-name* 例：	フィルタ ポリシーを有効にして、出力ロードバランシング（ELB）を対象の VTEP ルートのみに制限します。

	コマンドまたはアクション	目的
	`switch(config-router-af)# load-balance egress filter-policy route-map ROUTE_MAP_1` `switch(config-router-af)#`	コミュニティまたはプレフィックスリストでアンダーレイ ルートをタグ付けできます。 ポリシーをフィルタするには、このコマンドの **no** 形式を使用します。 （注） • アドバタイズする出力 BGW は、コミュニティでフィルタされる場合、同じコミュニティの指定のルートにすべてのパスをタグ付けする必要があります。 • フィルタ ポリシーが構成されていることを確認します。この構成を行わないと、システムはプレフィックスの ELB パスを計算しません。

アンダーレイの出力ロード バランス自動マルチパス ポリシーの作成

計算するアンダーレイ パスの最大数と、それらのアンダーレイ パスを同じマルチパス セットに割り当てる基準を指定するルートマップを使用して、自動マルチパス ポリシーを構成できます。このポリシーは、ベストパスと比較した場合に、AIGP メトリックや AS-Path の差など、1 つ以上の構成されたしきい値を照合できます。自動マルチパス ポリシーがない場合は、ベストパスのみがインストールされます。

アンダーレイの出力ロード バランシング自動マルチパス ポリシーを構成するには、次の手順を実行します。

手順の概要

1. **configure terminal**
2. **ip prefix-list** *prefix-list-name* **seq** *value* **permit** *nexthop-ipaddress*
3. **route-map** *route-map-name*
4. **exit**
5. **router bgp** *as-number*
6. **address-family ipv4 unicast**
7. **[no] load-balance egress multipath auto-policy route-map** *route-map-name*

手順の詳細

手順

ステップ 1 configure terminal

例：

```
switch# config terminal
switch(config)#
```

グローバル コンフィギュレーション モードを開始します。

（注）
プレフィックス リストが作成されていない場合にのみ、ステップ 2 に進みます。

ステップ 2 **ip prefix-list** *prefix-list-name* **seq** *value* **permit** *nexthop-ipaddress*

例：

```
switch(config)# ip prefix-list remote_nexthop seq 5 permit 10.10.112.1/32
```

リモート ネクストホップと一致するようにプレフィックス リストを構成します。

ステップ 3 **route-map** *route-map-name*

このルートマップは、アンダーレイ パスがベストパスに対して等しくない（下位の）場合でも、同じマルチパス セットの一部としてグループ化します。これは、ベストパスのこれらの値と比較したときに、これらのアンダーレイ パスの AIGP メトリックまたは AS パス長の構成された差に基づいて実行できます。

（注）
「match」および「set」コマンドを使用して、指定された要件に従ってシステムを構成します。

出力ロード バランシングの自動マルチパス ポリシーは、以下のように有効にできます：

例：

```
switch(config-route-map)# route-map ROUTE_MAP_2
```

a) **set maximum-paths** *max-path-value*

例：

```
switch(config-route-map)# set maximum-paths 5
```

出力ロードバランシングのために計算およびインストールされるマルチパスの最大数を構成します。範囲は、1 ～ 64 です。

および

b) **set as-path-length difference** *as-path-diff-value*

例：

```
switch(config-route-map)# set as-path-length difference 5
```

不等コスト ロードバランスに対してアンダーレイ パスで考慮する必要があるベストパスと比較して、AS-Path-length の差を構成します。範囲は 1 ～ 255 です。

c) **set aigp-metric difference** *value*

例：

```
switch(config-route-map)# set aigp-metric difference 100
```

不等コスト ロードバランスに対してアンダーレイ パスで考慮する必要があるベストパスと比較して、AIGP メトリック値の差を構成します。範囲は 1 ～ 4294967295 です。

（注）

AIGP メトリック情報を構成および使用する方法の詳細については、「VXLAN EVPN TE の構成例：マルチサイト出力ロードバランシング （18 ページ）」 を参照してください。

ステップ 4 exit

例：

```
switch(config-route-map)# exit
switch(config)#
```

BGP ルータ コンフィギュレーション モードを開始します。

ステップ 5 router bgp *as-number*

例：

```
switch(config)# router bgp 65001
switch(config-router)#
```

BGP ルータ コンフィギュレーション モードを開始します。

ステップ 6 address-family ipv4 unicast

例：

```
switch(config-router)# address-family ipv4 unicast
switch(config-router-af)#
```

IPv4 ユニ キャスト アドレス ファミリを構成します。

ステップ 7 [no] load-balance egress multipath auto-policy route-map *route-map-name*

例：

```
switch(config-router-af)# load-balance egress multipath auto-policy route-map ROUTE_MAP_2
```

BGP での自動マルチパス選択とロード シェアリングを制御するパラメータ構成します。

パラメータ構成を削除するには、このコマンドの **no** 形式を使用します。

自動マルチパス ポリシーがない場合は、ベスト パスのみがインストールされます。

（注）
コミュニティの一致を使用してプレフィックスを選択する場合、このプレフィックスのすべてのパスは、BGP への発信中にプレフィックスをタグ付けすることによって、出力 BGW をアドバタイズすることによって同じコミュニティでタグ付けされる必要があります。

アンダーレイのダイナミックの導出

前のセクションで説明した構成では、不均等なアンダーレイ パスをベストパスと一緒に同じマルチパスセットに追加して、これらすべてのパス間でトラフィックの均等なロードバランシング（uECMP）を有効にすることができました。

ここでは、代わりに、同じマルチパス セットのアンダーレイ パス部分に重みをスタティックに割り当てて、オーバーレイ トラフィック フローをより適切に分散する方法について説明し

ます。これには 2 つのオプションがあります。1 つ目は、ベストパスに異なるスタティック重みを割り当て、ベストパスと等しくない（劣る）アンダーレイパスのセットに異なるスタティック重みを割り当てることです。は、特定の基準（AS パス長または AIGP メトリック）に基づいてマルチパスセットに追加されます。2 番目のオプションでは、特定のアンダーレイ パスにスタティックな重みを明示的に割り当てることができます。

ロード シェア重みの計算

自動マルチパス ポリシーに次のコマンドのいずれかまたは両方が含まれている場合、BGP はロードシェア重み計算（LSWC）を使用して重みを導出します。

- **set load-share multipath-equal-group**
- **set load-share multipath-unequal-group**

（注）

- AS パス長に関してベスト パスと同等のアンダーレイ パスのセットは、 **multipath-equal-group** として分類されます。
- ベストパスの品質には一致しないが、AS パス長の指定された差分しきい値内にあるアンダーレイ パスのセットは、**multipath-unequal-group**として分類されます。

次の例 1 では、BGP は LSWC を使用して複数のアンダーレイ パスの重みを取得します。等しくないアンダーレイ パスは、最適パスと比較したときに AS パス長が構成の差（4）を超えている場合、 **multipath-unequal-group** に追加されます。

```
route-map auto-multipath  permit 10
  match ip address prefix-list site_A_BGW2
  set as-path length difference 4
  set maximum-paths 64
  set load-share multipath-equal-group 40
  set load-share multipath-unequal-group 60
```

次の例 2 では、最適パス AIGP メトリックと比較したときに、AIGP メトリックが構成の差（10）に収まる場合に、等しくないアンダーレイ パスの BGP が代わりに **multipath-unequal-group** に追加されます。

```
route-map auto-multipath  permit 10
  match ip address prefix-list site_A_BGW2
  set aigp-metric difference 10
  set maximum-paths 64
  set load-share multipath-equal-group 40
  set load-share multipath-unequal-group 60
```

前の例では、マルチパス グループの 1 つ（**load-share multipath-equal-group** または **load-share multipath-unequal-group**）のみが構成されている場合、もう一方は BGP によって値 1 に構成されていると暗黙的に想定されます。

明示的な負荷分散重みの計算

明示的な負荷分散パスを使用するには、宛先IPアドレスに関連付けられた特定のネクストホップまたはパスに一致するルートマップを定義する必要があります。この構成には、次の特性があります。

- これらは、auto-multipath route-mapエントリよりも優先される特定のroute-mapエントリです。
- 自動マルチパスルールは、明示的な負荷分散ルールとともに存在する場合と存在しない場合があります。
- 宛先ごとに許可される明示パス エントリの最大数は64です。これは、auto-multipathで指定された最大パス属性とは無関係です。
- 特定の宛先のパスと一致する明示的な負荷分散ルールの場合、そのパスは自動マルチパスルールによって含まれたり処理されたりしません（存在する場合）。これは、ECMPまたはuECMPとして認定されないパスである可能性もあります。
- 自動マルチパス（単独）では、最適パスである**multipath-equal-group**に適合するパスが常に存在します。宛先のベスト パスに一致する明示的な負荷分散ルールがある場合、**multipath-equal-group**のパスが0になる可能性があります。（これは、その宛先への他のECMPパスがない場合に発生します）。
- パスの明示的な負荷分散値は、自動マルチパス（存在する場合）で指定されたECMPまたはuECMP負荷分散値に比例します。
- 自動マルチパスでは、ECMPまたはuECMPグループ内のパスごとに個別の重みを生成するための計算が実行されます。この値は、各グループ内のパスの数と指定された負荷分散値の関数です。
- 明示的な負荷分散パスのみが定義されている場合（自動マルチパスなし）、コマンドで指定された値が負荷分散の重みとして使用されます。
- 明示的な負荷分散重みポリシーに指定された負荷分散がない場合、デフォルト値はなく、ポリシーは効果がありません。
- 「ベスト パス」として識別されるパスには常に重みが割り当てられ、URIBにダウンロードされます。これは、ベストパスに明示的な負荷分散ポリシーが指定されていない場合でも発生します。ベスト パスに明示的な負荷分散が指定されておらず、自動マルチパス ポリシーが存在しない場合、ベストパスはURIBに重み1で表示されます。

ロードバランシング出力明示ポリシーを有効にするには、次の手順を実行します。

手順の概要

1. **configure terminal**
2. **router bgp** *as-number*
3. **route-map** *routemap-name* **permit** *value*
4. **match ip address prefix-list** *dest-IPaddress*
5. **match ip next-hop prefix-list path** *value*

6. **set load-share** *value*

手順の詳細

手順

	コマンドまたはアクション	目的

ステップ 1	**configure terminal**
	例：
	`switch# config terminal`
	`switch(config)#`
		グローバル コンフィギュレーション モードを開始します。

ステップ 2	**router bgp** *as-number*
	例：
	`switch(config)# router bgp 100`
	`switch(config-router)#`
		BGP ルータ コンフィギュレーション モードを開始します。

ステップ 3	**route-map** *routemap-name* **permit** *value* 例： `switch(config)# route-map load_distribute permit 6`	ルートマップを作成するか、ルートマップ構成モードを開始します。
ステップ 4	**match ip address prefix-list** *dest-IPaddress* 例： `switch(config-route-map)# match ip address prefix-list match_ip`	アンダーレイ ルートは、プレフィックス リストを使用して一致させます。
ステップ 5	**match ip next-hop prefix-list path** *value* 例： `switch(config-route-map)# match ip next-hop prefix-list path1`	明示的な負荷分散パスは、宛先 IP アドレスに関連付けられた特定のネクストホップまたはパスと一致します。
ステップ 6	**set load-share** *value* 例： `switch(config-route-map)# set load-share 199`	コンフィギュレーションで指定された明示的なロード シェア パスを設定します。

例

set load-shareなどの明示的な負荷分散コマンドが、プレフィックスまたはネクストホップの一致に基づいて自動マルチパス ポリシー内で構成されている場合、ポリシーに 1 つ以上の LSWC マルチパスグループ コマンドのいずれかが含まれている場合にのみ有効になります。

次の例 1 では、 **set aigp-metric difference** が定義されていますが、BGP は LSWC を使用して重みを取得します。この場合、一致するプレフィックスとネクストホップに対して明示的な負荷分散が計算されます。

```
route-map auto-multipath permit 6
  match ip address prefix-list match_ip
  match ip next-hop prefix-list path1
  set load-share 100

route-map auto-multipath permit 10
   match ip address prefix-list site_A_BGW2
   set as-path-length difference 4
   set aigp-metric difference 100
   set maximum-paths 64
   set load-share multipath-equal-group 40
   set load-share multipath-unequal-group 60
```

次の例 2 では、BGP は AMWC を使用して重みを計算し、明示的な負荷分散は無視されます。

```
route-map auto-multipath permit 6
  match ip address prefix-list match_ip
  match ip next-hop prefix-list path1
  set load-share 100

route-map auto-multipath permit 10
   match ip address prefix-list site_A_BGW2
   set as-path-length difference 4
   set aigp-metric difference 100
   set maximum-paths 64
```

オーバーレイの出力ロード バランシングの有効化

アンダーレイで **egress-loadbalance-resolution-vrf** を使用してオーバーレイ（EVPN）プレフィックスのネクストホップ解決を有効にするには、次の手順を実行します。

手順の概要

1. **configure terminal**
2. **router bgp** *as-number*
3. **address-family l2vpn evpn**
4. **[no] nexthop load-balance egress multisite**

手順の詳細

手順

	コマンドまたはアクション	目的

ステップ 1	**configure terminal**
	例：
	`switch# config terminal`
	`switch(config)#`
		グローバル コンフィギュレーション モードを開始します。

	コマンドまたはアクション	目的

ステップ 2	**router bgp** *as-number*
	例：
	`switch(config)# router bgp 100`
	`switch(config-router)#`
		BGP ルータ コンフィギュレーション モードを開始します。

ステップ 3	**address-family l2vpn evpn**
	例：
	`switch(config-router)# address-family l2vpn evpn`
	`switch((config-router-af)#`
		L2VPN アドレス ファミリを設定します。

ステップ 4	**[no] nexthop load-balance egress multisite** 例： `switch(config-router-af)# nexthop load-balance egress multisite`	**egress-loadbalance-resolution- VRF** の対応する IPv4 または IPv6 テーブルを使用して、オーバーレイ（EVPN）ネクストホップ解決を有効にします。出力ロードバランス解決 VRF のテーブルの詳細については、 VXLAN EVPN TE：マルチサイト出力ロードバランシングの構成 （4 ページ） を参照してください。 オーバーレイの出力ロードバランシング構成を削除するには、このコマンドの **no** 形式を使用します。 **multisite** オプションは、この機能をマルチサイトネットワークから学習された EVPN ネクストホップのみに制限します。これは、さまざまな VRF にインポートされたタイプ 2 とタイプ 5 の両方のルートに適用されます。 （注） この構成は、アンダーレイ テーブルで出力ロードバランスの計算を有効にした後に有効にする必要があります。

オーバーレイの uECMP または wuECMP ロードバランシングの有効化

複数のオーバーレイ ネクストホップ間で不等コストまたは重み付けロードバランシングを有効にできます。

（注） この構成では、VIP ネクストホップはサポートされません。

手順

	コマンドまたはアクション	目的

ステップ 1	**configure terminal**
	例：
	`switch# config terminal`
	`switch(config)#`
		グローバル コンフィギュレーション モードを開始します。

ステップ 2	**router bgp** *as-number*
	例：
	`switch(config)# router bgp 100`
	`switch(config-router)#`
		BGP ルータ コンフィギュレーション モードを開始します。

ステップ 3	**address-family l2vpn evpn**
	例：
	`switch(config-router)# address-family l2vpn evpn`
	`switch((config-router-af)#`
		L2VPN アドレス ファミリを設定します。

ステップ 4	**nexthop load-balance egress multisite** 例： `switch(config-router-af)# nexthop load-balance egress multisite`	**egress-loadbalance-resolution- VRF** の対応する IPv4 または IPv6 テーブルを使用して、オーバーレイ（EVPN）ネクストホップ解決を有効にします。**egress-loadbalance-resolution- VRF's** のテーブルの詳細については、VXLAN EVPN TE：マルチサイト出力ロードバランシングの構成 （4 ページ） を参照してください。
ステップ 5	**[no] maximum-path** *value* 例： `switch(config-router-af)# maximum-path 64`	出力ロードバランシングに指定された最大パスを構成します。 出力ロード バランシングの最大パスを削除するには、このコマンドの **no** 形式を使用します。 （注） 指定された値が 1 より大きい場合、以下のコマンド（maximum-paths unequal-cost）を使用すると、オーバーレイ ネクストホップの wuECMP が有効になり、ネクストホップのネクストホップ メトリックに基づいて重みが自動的に導出されます。
ステップ 6	**[no] maximum-paths unequal-cost** 例： `switch(config-router-af)# maximum-path unequal-cost`	オーバーレイで不等マルチパスを構成します。 オーバーレイで不等マルチパスを無効にするには、このコマンドの **no** 形式を使用します。 （注） • **nexthop load-balance egress multisite** コマンドが **maximum-path**および **maximum-path unequal**

	コマンドまたはアクション	目的
		コマンドとともに構成されている場合、オーバーレイ ネクストホップは、複数のオーバーレイ ネクストホップがあり、それらのネクストホップの **igp_metrics** が異なる場合にのみ、重みでプログラムされます。オーバーレイ ネクストホップは、 **egress-loadbalance-resolution-VRF** テーブルを使用して解決されます。 • **nexthop load-balance egress multisite** コマンドが構成されていないが、 **maximum-path**、**maximum-path unequal** コマンドが構成されている場合、オーバーレイ ネクストホップは、複数のオーバーレイ ネクストホップがあり、それらのネクストホップの **igp_metrics** が異なる場合にのみ、重みでプログラムされます。オーバーレイのネクストホップは、 **デフォルト** のテーブルを使用して解決されます。

ステップ **7**	**exit**
	例：
	`switch(config-router-af)# exit`
	`switch(config-router)#`
		コマンド モードを終了します。

ステップ **8**	（任意） **[no] bestpath igp-metric ignore** 例： `switch(config-router)# bestpath igp-metric ignore`	この構成により、オーバーレイのネクストホップの igp_metric がベスト パスによって無視されるため、**maximum-paths unequal-cost** が構成されている場合でもオーバーレイに（wuECMP からの）ECMP が構成されます。 ECMP/wuECMP を無効にするには、このコマンドの **no** 形式を使用します。

VXLAN EVPN TE の確認：マルチサイト出力ロード バランシング構成の確認

VXLAN EVPN TE：マルチサイト出力ロードバランシングの構成情報を表示するには、次のいずれかのコマンドを入力します。

コマンド	目的
show ip|ipv6 route [detail] vrf egress-loadbalance-resolution-	自動的に作成される特別な内部 VRF を表示します。この VRF は、BGP で出力ロード バランス構成が有効になっているときに、内部で暗黙的に使用されます。 BGP が ELB フィルタポリシーと自動マルチパス ポリシーで構成されている場合、デフォルト テーブルからルートのベスト パスを継承し、ELB ポリシーに基づいて追加の ELB パスを含めます。 **detail** オプションが有効になっている場合、wuECMP が構成されている場合、BGP が RIB に送信する重みが表示されます。 （注） Cisco NX-OS リリース 10.4(3)F 以降、**egress-loadbalance-resolution-** VRF のテーブル ID は、値 4089/0x0ff9 でスタティックに割り当てられ、割り当ての「limit-resource vrf」プールの外になります。 既存のユーザー構成には影響しません。
show ip|ipv6 route [detail] vrf tenant_vrf	デフォルト テーブルの代わりに **egress-loadbalance-resolution-table** を介してネクストホップが解決される EVPN-VXLAN テナント VRF のオーバーレイ プレフィックスを表示します。 **detail** オプションを有効にすると、wuECMP が構成されている場合、BGP から RIB に送信されるネクストホップに割り当てられた重みが表示されます。
show bgp ipv4 unicast *ipaddress* **vrf egress-loadbalance-resolution-**	AIGP がアンダーレイで構成されている場合は、派生 AIGP メトリックを含む、アンダーレイ BGP ルートとネクストホップを表示します。wuECMP の場合、ダイナミックに（つまり、AIGP メトリックまたはスタティックに構成された負荷分散重みから）導出される重みが表示されます。uECMP または ECMP の場合、重みは表示されません。
show l2route evpn mac all detail	wuECMP が構成されている場合、MAC ルートのネクストホップと重みを表示します。

コマンド	目的
show l2route evpn ead es detail	wuECMP が構成されている場合、EAD/ES ルートのネクストホップに関連付けられた重みを表示します。

VXLAN EVPN TE の構成例：マルチサイト出力ロードバランシング

このセクションでは、次の使用例での VXLAN EVPN TE - マルチサイト出力ロードバランシングの構成と検証の詳細について説明します。

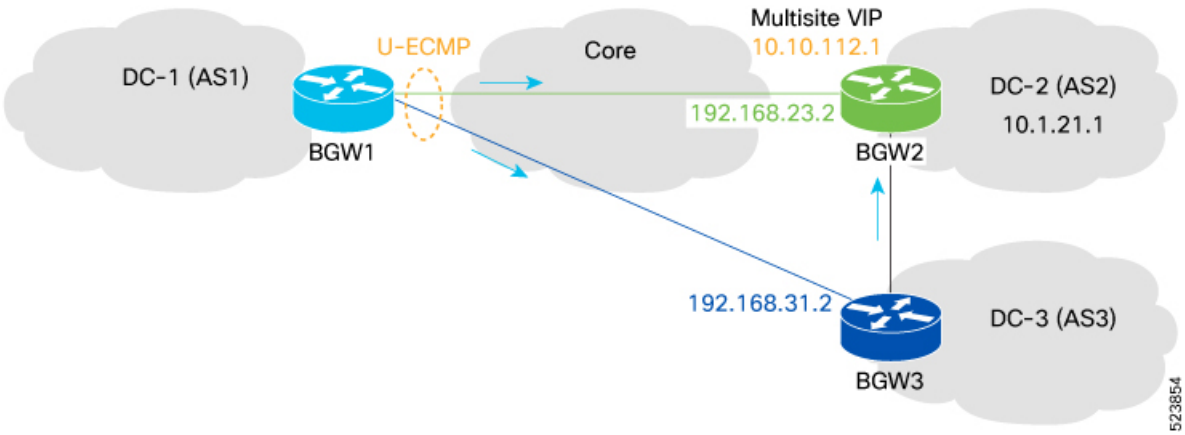
- アンダーレイの uECMP、オーバーレイ プレフィックスの単一の EVPN ネクストホップ。
- アンダーレイでのロードシェアリングと明示的なロードシェアリングを使用した静的 wuECMP、およびオーバーレイ プレフィックスの単一 EVPN ネクストホップ
- オーバーレイとアンダーレイのダイナミック重み（wuECMP）、およびアンダーレイの AIGP

アンダーレイの uECMP、オーバーレイ プレフィックスの単一の EVPN ネクストホップ

次の図で強調表示されているように、この使用例では、DC-2 から DC-1 にアドバタイズされるすべてのオーバーレイ プレフィックスには、DC-2 のすべての BGW デバイスによって共有されるマルチサイト VIP アドレスで表される単一の EVPN ネクストホップがあります（わかりやすくするために、1 つのデバイス（BGW2）のみが図に示されています）。

DC-1 の BGW1 デバイスから DC-2 のマルチサイト VIP へのアンダーレイ到達可能性は、次の 2 つのアンダーレイ パスを使用して可能です。

1. 緑色のパス：BGW1 を BGW2 に接続するベスト パス
2. 青色のパス：DC-3 の BGW3 を通過する好ましくないパス（または、最適ではない青色のパスは、コア インフラストラクチャの 1 つ以上のルータ デバイス部分を通過する可能性があります）。



この使用例の目的は、DC-1 と DC2 間のオーバーレイ通信に、緑と青の両方のアンダーレイ パスを均等に使用できるようにすることです。これを可能にするには、2 つのパスを同じマルチパス セットの uECMP パスの一部と見なす必要があり、DC-1 の BGW1 デバイスに適用される次の構成手順でこの目的を達成できます。

アンダーレイでの VXLAN EVPN TE - マルチサイト出力ロードバランシング uECMP の有効化

次のすべてのコマンドは、DC-1 の BGW1 デバイスで構成する必要があります。

1. フィルタ ポリシーを作成します。
   - アンダーレイ ルートを指定して、ELB uECMP を有効にします。この場合、アンダーレイ ルートは DC-2 のマルチサイト VIP アドレスであり、DC-1 にアドバタイズされたオーバーレイ プレフィックスのネクストホップを示します。

     ```
     ip prefix-list site2_ms_vip seq 5 permit 10.10.112.1/32
     ```

   - ルートマップを作成し、照合状態で以前構成されたプレフィックスリストを適用します。

     ```
     route-map Filter-Policy permit 10
       match ip address prefix-list site2_ms_vip
     ```

2. ELB フィルタポリシーを有効にします（IPv4 または IPv4 BGP アドレスファミリ下）。

   ```
   router bgp 1
     address-family ipv4 unicast
       load-balance egress filter-policy route-map Filter-Policy
   ```

3. **Egress-loadbalance-resolution-**VRF テーブルで DC-2 マルチサイト VIP ルートのインストールを確認します。この時点では、緑色のベスト パスのみが宛先に到達すると見なされます。

   ```
   BGW1# show ip route 10.10.112.1 vrf egress-loadbalance-resolution-

   10.10.112.1/32, ubest/mbest: 1/0
     *via 192.168.23.2%default, [20/0], 16:36:15, bgp-1, external, tag 2, uecmp ! Green
    path
   ```

4. マルチパス自動ポリシーを作成します。

- 不均等なアンダーレイ パスをベスト パスとともにマルチパス セットに割り当てる基準を指定します。AS パス長などの BGP 属性の違いは、同じマルチパス セットの一部として不均等なアンダーレイ パスをグループ化すると見なすことができます。
- 「出力ロードバランシング」のために計算され、インストールされるマルチパスセットの一部であるアンダーレイ パスの最大数を指定します。

```
route-map Auto-Policy permit 10
  set as-path-length difference 1 <1 to 255>
  set maximum-paths 2 <1 to 64>
```

5. IPv4 または IPv6 BGP アドレスファミリで ELB マルチパス自動ポリシーを有効にします。

```
router bgp 1
  address-family ipv4 unicast
    load-balance egress multipath auto-policy route-map Auto-Policy
```

6. URIB の Egress-loadbalance-resolution VRF テーブルのルートを確認します。DC-2 のマルチサイト VIP アドレスに到達するための実行可能なオプションとして、青色の最適ではないパスが緑色のベスト パスに追加されました。

```
BGW1# show ip route 10.10.112.1 vrf egress-loadbalance-resolution-

10.10.112.1/32, ubest/mbest: 2/0
  *via 192.168.23.2%default, [20/0], 00:44:17, bgp-1, external, tag 2, uecmp ! Green
 path
  *via 192.168.31.2%default, [20/0], 00:44:17, bgp-1, external, tag 3, uecmp ! Blue
 path
```

オーバーレイでの uECMP の ELB VRF での解決の有効化

1. 出力ロードバランシング uECMP アンダーレイ パスを使用したオーバーレイ ネクストホップ解決を有効にします。
   - このコマンドは、**Egress-loadbalance-resolution-** VRF テーブルを使用した EVPN ネクストホップの解決を有効にします。したがって、両方のアンダーレイ パス（緑と青の等しくないパス）は、サイト間トラフィックの等コスト ロードバランシングに使用されます。
   - このコマンドは、アンダーレイ テーブルで出力ロード バランシング計算を有効にした後、BGP L2VPN EVPN アドレス ファミリで有効にする必要があります。

```
router bgp 1
  address-family l2vpn evpn
    nexthop load-balance egress multisite
```

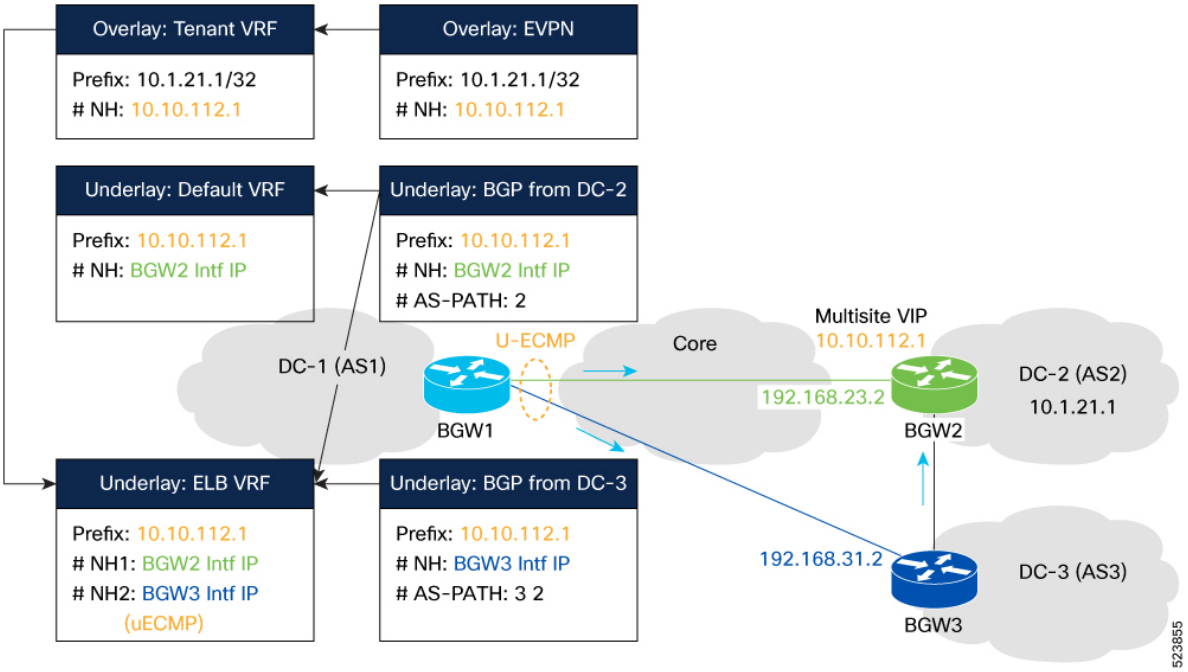
2. テナント VRF テーブルでオーバーレイ ルートを確認します。次に示すように、DC-1 の BGW1 デバイスから DC-2 のオーバーレイ プレフィックスへの到達可能性は、DC-2 マルチサイト VIP アドレス（10.10.112.1）を介して行われ、そのルックアップは **egress-loadbalance-resolution-**で実行されます。その結果、トラフィックはインストールされた 2 つのアンダーレイ パスに均等に分散されます。（上記参照）。

```
BGW1# show ip route 10.1.21.1 vrf 3001

10.1.21.1/32, ubest/mbest: 1/0
  *via 10.10.112.1%egress-loadbalance-resolution-, [20/2000], 04:43:40, bgp-1,
```

```
external, tag 2,
  eLB, segid: 3003001 tunnelid: 0x67027001 encap: VXLAN
```

uECMP ルーティングの詳細

このセクションでは、アンダーレイの uECMP、オーバーレイ プレフィックスの単一 EVPN ネクストホップの使用例に関する詳細なトラブルシューティング情報を提供します。



1. **アンダーレイ ルート プログラミング**

   uECMP アンダーレイ パスを介してリモート BGW から受信した DC-2 マルチサイト VIP ルートは、DC-1 の BGW1 デバイスで、出力ロードバランス解決 VRF の BGP ルーティング テーブルとユニキャスト ルーティング テーブルにプログラムされます。

   以下の例では、以下のコンポーネントの出力例を示します。

   - BGP - uECMP：AS-Path の差が 1 であるため、2 番目のパスが **multipath uecmp** としてラベル付けされており、以前定義したアンダーレイ パスを同じマルチパス セットの一部とみなす基準と一致していることに注意してください。

```
BGW1# show bgp ipv4 unicast 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.23.2 (metric 0) from 192.168.23.2 (101.2.33.1)
      Origin incomplete, MED 0, localpref 100, weight 0

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
```

```
    192.168.31.2 (metric 0) from 192.168.31.2 (101.3.33.1)
      Origin incomplete, MED not set, localpref 100, weight 0
```

- URIB - uECMP：次の例は、2 つの uECMP パスが DC-2 マルチサイト VIP 宛先に到達するために均等に活用される方法を示しています。

```
BGW1# show ip route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
10.10.112.1/32, ubest/mbest: 2/0
    *via 192.168.23.2%default, [20/0], 17:58:44, bgp-1, external, tag 2, uecmp
! Green path
    *via 192.168.31.2%default, [20/0], 17:58:44, bgp-1, external, tag 3, uecmp
! Green path
```

- FIB - uECMP：

```
BGW1# show forwarding route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
*10.10.112.1/32      192.168.23.2           Ethernet1/6
                     192.168.31.2           Ethernet1/7
```

2. オーバーレイ ルート プログラミング

   egress-loadbalance-resolution- VRF は、受信したオーバーレイ ルートに関連付けられたリモート VIP ネクストホップを解決するために使用されます。

   以下の例では、以下のコンポーネントの出力例を示します。

   - BGP：

```
BGW1# show bgp l2vpn evpn 10.1.21.1
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, has esi_gw

             Imported to 3 destination(s)
             Imported paths list: 3001 L3-3003001 L2-2001001
  AS-Path: 2 , path sourced external to AS
    10.10.112.1 (metric 0) from 101.2.33.1 (101.2.33.1)
      Origin IGP, MED 2000, localpref 100, weight 0
      Received label 2001001 3003001
      Extcommunity: RT:1:2001001 RT:1:3003001 SOO:102.2.121.1:0
```

   - URIB：

```
BGW1# show ip route 10.1.21.1 vrf 3001
<Truncated>
10.1.21.1/32, ubest/mbest: 1/0
    *via 10.10.112.1%egress-loadbalance-resolution-, [20/2000], 00:28:03, bgp-1,
 external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x67027001 encap: VXLAN
```

   - FIB：

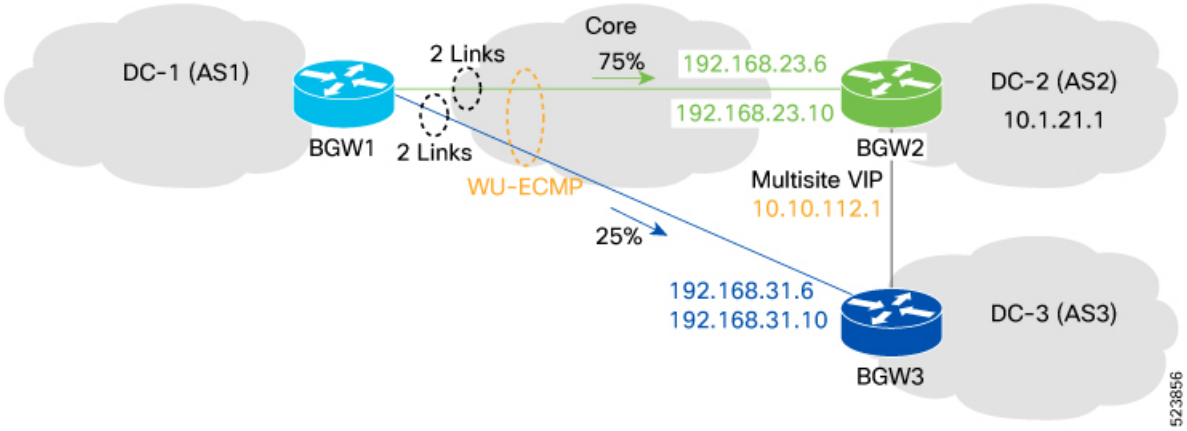
```
BGW1# show forwarding route 10.1.21.1 vrf 3001
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
```

```
-----------------+--------------------+---------------+----------------+---------------
10.1.21.1/32         10.10.112.1           nve1

BGW1# show forwarding route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix             | Next-hop             | Interface       | Labels           |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
*10.10.112.1/32      192.168.23.2            Ethernet1/6
                     192.168.31.2            Ethernet1/7
```

アンダーレイでの負荷分散と明示的な負荷分散を使用したスタティック wuECMP、およびオーバーレイ プレフィックスの単一 EVPN ネクストホップ

アンダーレイでの負荷分散を使用した VXLAN EVPN TE - マルチサイト出力ロードバランシング静的 wuECMP の有効化

次の図で強調表示されているように、この使用例でも、DC-2 から DC-1 にアドバタイズされるすべてのオーバーレイ プレフィックスには、DC-2 のすべての BGW デバイスによって共有されるマルチサイト VIP アドレスで表される単一の EVPN ネクストホップがあります。 DC-1 の BGW1 デバイスから DC-2 の Multi-Site VIP へのアンダーレイ到達可能性は、4 つのアンダーレイパスを介して可能です。 BGW1 と BGW2 を結ぶ 2 つの緑色の ECMP ベスト パスと、DC- 3 の BGW3 を経由する 2 つの不利なアンダーレイ パス （または、最適ではない青色のパスは、コアインフラストラクチャの 1 つ以上のルータ デバイス部分を通過する可能性があります）。



この使用例の目的は、DC-1 と DC2 間のオーバーレイ通信に緑色と青色の両方のアンダーレイパスを使用できるようにすることですが、トラフィックの負荷分散は異なります（トラフィックの 75% は緑色のパスを使用し、青色のパスの 25%）。DC-1 の BGW1 デバイスに適用される次の構成手順は、この目的を実現します。

1. フィルタポリシーを作成します。

- アンダーレイ ルートを指定して、ELB wuECMP を有効にします。 この場合、アンダーレイ ルートは DC-2 のマルチサイト VIP アドレスであり、DC-1 にアドバタイズされたオーバーレイ プレフィックスのネクストホップを示します。

```
ip prefix-list site2_ms_vip seq 5 permit 10.10.112.1/32
```

- ルートマップを作成し、照合状態で以前構成されたプレフィックスリストを適用します。

```
route-map Filter-Policy permit 10
  match ip address prefix-list site2_ms_vip
```

2. ELB フィルタポリシーを有効にします（IPv4 または IPv4 BGP アドレスファミリ下）。

```
router bgp 1
  address-family ipv4 unicast
    load-balance egress filter-policy route-map Filter-Policy
```

3. **Egress-loadbalance-resolution-** VRF テーブルのルートを確認します。デフォルトでは、DC-2 のマルチサイト VIP アドレスに到達するために、2 つの緑色の ECMP アンダーレイパスのみがインストールされます。

```
BGW1# show ip route 10.10.112.1 vrf egress-loadbalance-resolution-

10.10.112.1/32, ubest/mbest: 1/0
  *via 192.168.23.6%default, [20/0], 16:36:15, bgp-1, external, tag 2, uecmp ! Green
 Path
  *via 192.168.23.10%default, [20/0], 16:37:41, bgp-1, external, tag 2, uecmp ! Green
 Path
```

4. マルチパス自動ポリシーを作成します。詳細については、「ロードシェア重みの計算（10 ページ）」を参照してください。

- 同じマルチパス セットの一部としてアンダーレイ パスをグループ化すると見なすことができる AS-Path 長などの BGP 属性の違いを指定します。
- 「出力ロードバランシング」のために計算され、インストールされるマルチパスの最大数を指定します。この場合、値は 4（2 つの緑と 2 つの青のパス）です。
- アンダーレイパスの ECMP パスセットおよび uECMP パスセットに関連付ける負荷分散の重みを指定します。目標は、上の図に示すように、トラフィックの 75% を緑色のベスト パス アンダーレイ リンクで送信し、25% を青色の等しくないアンダーレイ リンクで送信することです。

```
route-map Auto-Policy permit 10
  set as-path-length difference 1 <1 to 255>
  set maximum-paths 4 <1 to 64>
  set load-share multipath-equal-group 3 <1 to 255>
  set load-share multipath-unequal-group 1 <1 to 255>
```

5. IPv4 または IPv6 BGP アドレスファミリで ELB マルチパス自動ポリシーを有効にします。

```
router bgp 1
  address-family ipv4 unicast
    load-balance egress multipath auto-policy route-map Auto-policy
```

6. URIB の **Egress-loadbalance-resolution** VRF テーブルで負荷共有ルートを使用した wuECMP を確認します。DC-2 のマルチサイト VIP アドレスに到達するための実行可能なオプショ

ンとして、青色の準最適パスが緑色のベストパスに追加されましたが、重みは異なります（1）。

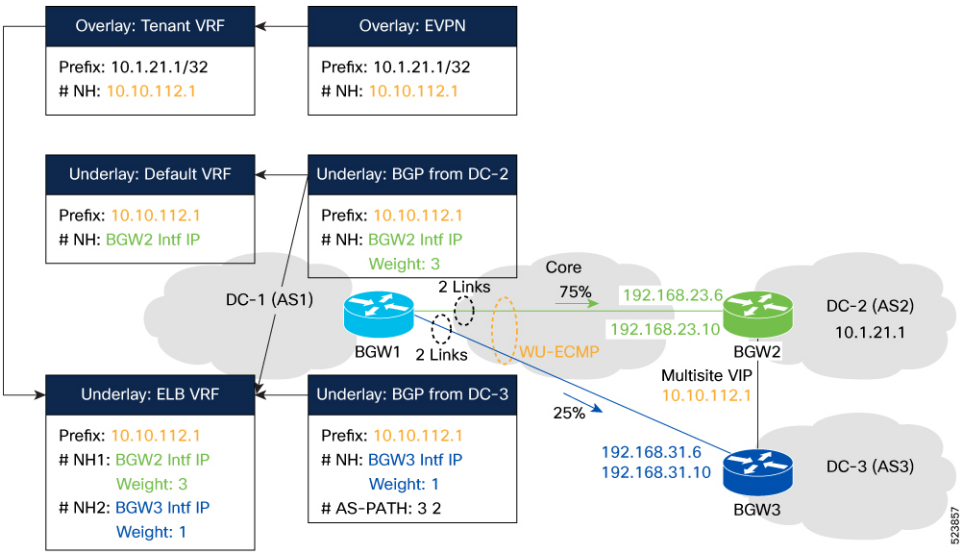
```
BGW1# show ip route 10.10.112.1 detail vrf egress-loadbalance-resolution-

10.10.112.1/32, ubest/mbest: 4/0
    *via 192.168.23.6%default, [20/0], 1w0d, bgp-1, weight:3, external, tag 2, uecmp
 ! Green Path
    *via 192.168.23.10%default, [20/0], 1w0d, bgp-1, weight:3, external, tag 2, uecmp
 ! Green Path
    *via 192.168.31.6%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3, uecmp
 ! Blue Path
    *via 192.168.31.10%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3, uecmp
 ! Blue Path
```

ECMP パスと uECMP パスに適用される自動負荷分散構成により、2 つの緑のパスには合計重み 3（それぞれ 1.5）が割り当てられ、2 つの青いパスには合計重み 1（それぞれ 0.5）が割り当てられます。その結果、DC-1 から DC-2 へのオーバーレイ トラフィックの 75% は緑色のパスを介してルーティングされ、残りの 25% は青色のパスを通過します。

緑色のパス：(3 + 3) / (3+3+1+1) = 0.75

青色のパス：(1 + 1) / (3+3+0.5+0.5) = 0.25

負荷分散ルーティングの詳細を使用した静的 wuECMP



1. **アンダーレイ ルート プログラミング**

uECMP アンダーレイ パスを介してリモート BGW から受信した DC-2 マルチサイト VIP ルートは、DC-1 の BGW1 デバイスで、**egress-loadbalance-resolution-** VRF の BGP ルーティング テーブルとユニキャスト ルーティング テーブルにプログラムされます。

以下の例では、以下のコンポーネントの出力例を示します。

- BGP：実際には最初のベストパス 1 を持つ ECMP パスであっても、2 番目のパスが「**multipath uecmp**」とラベル付けされていることに注意してください。AS-Path の差

は1であり、先に定義した条件に一致するアンダーレイパスを、同じマルチパスセットの一部とみなすため、他の2つのパスは適切な「**multipath uecmp**」です。さらに、負荷分散構成のため、3:1 の重み比が 2 つのパスのセットに適用されます。

```
BGW1# show bgp ipv4 unicast 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.23.6 (metric 0) from 192.168.23.6 (101.2.33.1)
      Origin incomplete, MED 0, localpref 100, weight 0, load share weight 3

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.23.10 (metric 0) from 192.168.23.10 (101.2.33.1)
      Origin incomplete, MED 0, localpref 100, weight 0, load share weight 3
  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
    192.168.31.6 (metric 0) from 192.168.31.6 (101.3.33.1)
     Origin incomplete, MED not set, localpref 100, weight 0, load share weight
 1

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
    192.168.31.10 (metric 0) from 192.168.31.10 (101.3.33.1)
      Origin incomplete, MED not set, localpref 100, weight 0, load share weight
 1
```

- URIB：

```
BGW1# show ip route 10.10.112.1 detail vrf egress-loadbalance-resolution-
<Truncated>
10.10.112.1/32, ubest/mbest: 4/0
    *via 192.168.23.6%default, [20/0], 1w0d, bgp-1, weight:3, external, tag 2,
uecmp ! Green Path
<Truncated>
    *via 192.168.23.10%default, [20/0], 1w0d, bgp-1, weight:3, external, tag 2,
 uecmp ! Green Path
<Truncated>
    *via 192.168.31.6%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3,
uecmp ! Blue Path
<Truncated>
    *via 192.168.31.10%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3,
 uecmp ! Blue Path
<Truncated>
```

- FIB：

```
BGW1# show forwarding route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
*10.10.112.1/32      192.168.23.6           Ethernet1/22 >> 24 Entries
                     192.168.23.6           Ethernet1/22
```

```
.....
192.168.23.10           Ethernet1/23 >> 24 Entries
192.168.23.10           Ethernet1/23
.....
192.168.31.6            Ethernet1/32 >> 8 Entries
192.168.31.6            Ethernet1/33
.....
192.168.31.10           Ethernet1/32 >> 8 Entries
192.168.31.10           Ethernet1/33
.....
```

次に、ECMP（3）およびuECMP（1）パスセットの重みに基づいて作成されたFIB転送エントリの概要を示します。

- ECMP パスセット：48 エントリ
- uECMP パスセット：16 エントリ
- トラフィック比率は 48 : 16 = 3 : 1

2. **オーバーレイ ルート プログラミング**

egress-loadbalance-resolution-VRF テーブルは、DC-2 から DC-1 にアドバタイズされるオーバーレイ プレフィックスの VIP ネクストホップを解決するために使用されます。

以下の例では、以下のコンポーネントの出力例を示します。

- BGP：

```
BGW1# show bgp l2vpn evpn 10.1.21.1
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, has esi_gw

             Imported to 3 destination(s)
             Imported paths list: 3001 L3-3003001 L2-2001001
  AS-Path: 2 , path sourced external to AS
    10.10.112.1 (metric 0) from 101.2.33.1 (101.2.33.1)
      Origin IGP, MED 2000, localpref 100, weight 0
      Received label 2001001 3003001
      Extcommunity: RT:1:2001001 RT:1:3003001 SOO:102.2.121.1:0
```

- URIB：

```
BGW1# show ip route 10.1.21.1 vrf 3001
<Truncated>
10.1.21.1/32, ubest/mbest: 1/0
    *via 10.10.112.1%egress-loadbalance-resolution-, [20/2000], 00:28:03, bgp-1,
 external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x67027001 encap: VXLAN
```

- FIB：

```
BGW1# show forwarding route 10.1.21.1 vrf 3001
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
10.1.21.1/32        10.10.112.1          nve1
```

```
BGW1# show forwarding route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
*10.10.112.1/32      192.168.23.6           Ethernet1/22 >> 24 Entries
                     192.168.23.6           Ethernet1/22
                     .....
                     192.168.23.10          Ethernet1/23 >> 24 Entries
                     192.168.23.10          Ethernet1/23
                     .....
                     192.168.31.6           Ethernet1/32 >> 8 Entries
                     192.168.31.6           Ethernet1/33
                     .....
                     192.168.31.10          Ethernet1/32 >> 8 Entries
                     192.168.31.10          Ethernet1/33
                     .....
```

次に、ECMP（3）およびuECMP（1）パスセットの重みに基づいて作成されたFIB転送エントリの概要を示します。

- ECMP パスセット：48 エントリ
- uECMP パスセット：16 エントリ
- トラフィック比率は 48 : 16 = 3 : 1 です。

オーバーレイでの負荷分散を使用した静的 wuECMP の ELB VRF での解決の有効化

上記の両方の使用例では、マルチサイト VIP ネクストホップアドレスの解決が **Egress-loadbalance-resolution-** VRF テーブルで行われるようにする必要があります。

1. 出力ロードバランシング uECMP アンダーレイ パスを使用したオーバーレイ ネクストホップ解決を有効にします。
   - このコマンドは、**Egress-loadbalance-resolution-** VRF テーブルを使用した EVPN ネクストホップの解決を有効にします。
   - アンダーレイ テーブルで出力ロード バランシング計算を有効にした後、このコマンドを有効にする必要があります。

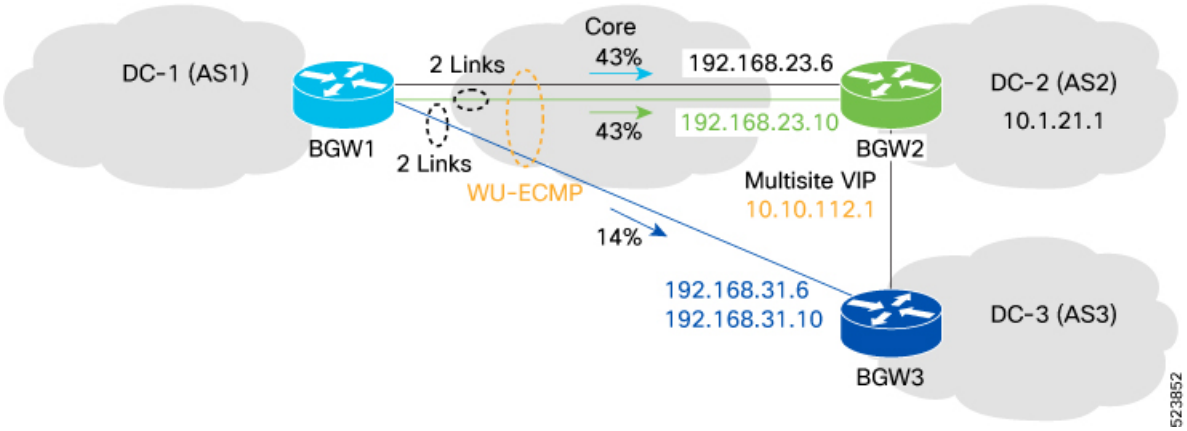
   ```
   router bgp 1
     address-family l2vpn evpn
       nexthop load-balance egress multisite
   ```

2. テナント VRF テーブルでオーバーレイ ルートを確認します。

   ```
   BGW1# show ip route 10.1.21.1 vrf 3001

   10.1.21.1/32, ubest/mbest: 1/0
     *via 10.10.112.1%egress-loadbalance-resolution-, [20/2000], 04:43:40, bgp-1,
   external, tag 2,
     eLB, segid: 3003001 tunnelid: 0x67027001 encap: VXLAN
   ```

アンダーレイでの明示的な負荷分散を使用した VXLAN EVPN TE - マルチサイト出力ロードバランシング静的 wuECMP の有効化

次の図は、前述の使用例をわずかに変更したものです。この使用例では、明示的な負荷分散構成を 2 つの緑色のパスのいずれかに適用できます。これにより、特定の緑色のパスが **multipath-equal-group** セットから除外され、緑色のパス間のトラフィック負荷の全体的な分散が変更されます。



次の構成手順は BGW1 に適用され、前の使用例で説明した手順と非常によく似ていますが、明示的な負荷分散構成が追加されているだけです。

1. フィルタポリシーを作成します。
   - アンダーレイ ルートを指定して、ELB wuECMP を有効にします。 またこの場合、アンダーレイルートは、DC-1 にアドバタイズされるオーバーレイ プレフィックスのネクストホップを表す DC-2 のマルチサイト VIP アドレスです。

     ```
     ip prefix-list site2_ms_vip seq 5 permit 10.10.112.1/32
     ```

   - ルートマップを作成し、照合状態で以前構成されたプレフィックスリストを適用します。

     ```
     route-map Filter-Policy permit 10
       match ip address prefix-list site2_ms_vip
     ```

2. ELB フィルタポリシーを有効にします（IPv4 または IPv4 BGP アドレスファミリ下）。

   ```
   router bgp 1
     address-family ipv4 unicast
       load-balance egress filter-policy route-map Filter-Policy
   ```

3. **Egress-loadbalance-resolution-** VRF テーブルのルートを確認します。

   ```
   BGW1# show ip route 10.10.112.1 vrf egress-loadbalance-resolution-

   10.10.112.1/32, ubest/mbest: 2/0
     *via 192.168.23.6%default, [20/0], 16:36:15, bgp-1, external, tag 2, uecmp ! Green
    Path
     *via 192.168.23.10%default, [20/0], 16:37:41, bgp-1, external, tag 2, uecmp ! Green
    Path
   ```

4. マルチパス自動ポリシーを作成します。詳細については、「ロードシェア重みの計算（10 ページ）」を参照してください。

   - マルチパスを選択するときに考慮することができる AS パス長などの BGP 属性の違いを指定します。
   - 「出力ロードバランシング」のために計算され、インストールされるマルチパスの最大数を指定します。
   - **明示的負荷分散**の場合、オーバーレイ ネクストホップに関連付けられた特定のネクストホップに一致する明示パスに負荷分散重みを割り当てることができます（この特定の例では、ネクストホップが 192.168.23.6 の緑色のパス）。

```
ip prefix-list MS_NH_S2_1 seq 5 permit 192.168.23.6/32

route-map Auto-Policy permit 5
  match ip address prefix-list site2_ms_vip
  match ip next-hop prefix-list MS_NH_S2_1
  set load-share 3 <1 to 255>

route-map Auto-Policy permit 10
   set as-path-length difference 1 <1 to 255>
   set maximum-paths 4 <1 to 64>
   set load-share multipath-equal-group 3 <1 to 255>
   set load-share multipath-unequal-group 1 <1 to 255>
```

5. ELB マルチパス自動ポリシーを有効にします。

```
router bgp 1
  address-family ipv4 unicast
    load-balance egress multipath auto-policy route-map Auto-policy
```

6. URIB の **Egress-loadbalance-resolution**VRF テーブルで明示的な負荷共有ルートを使用して wuECMP を確認します。DC-2 のマルチサイト VIP アドレスに到達するための実行可能なオプションとして、青色の準ベスト パスが緑色のベストパスに追加されました。

```
BGW1# show ip route 10.10.112.1 detail vrf egress-loadbalance-resolution-
<Truncated>
10.10.112.1/32, ubest/mbest: 4/0
    *via 192.168.23.6%default, [20/0], 1w0d, bgp-1, weight:6, external, tag 2, uecmp
 ! Green Path
<Truncated>
    *via 192.168.23.10%default, [20/0], 1w0d, bgp-1, weight:6, external, tag 2, uecmp
 ! Green Path
<Truncated>
    *via 192.168.31.6%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3, uecmp
 ! Blue Path
<Truncated>
    *via 192.168.31.10%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3, uecmp
 ! Blue Path
<Truncated>
```

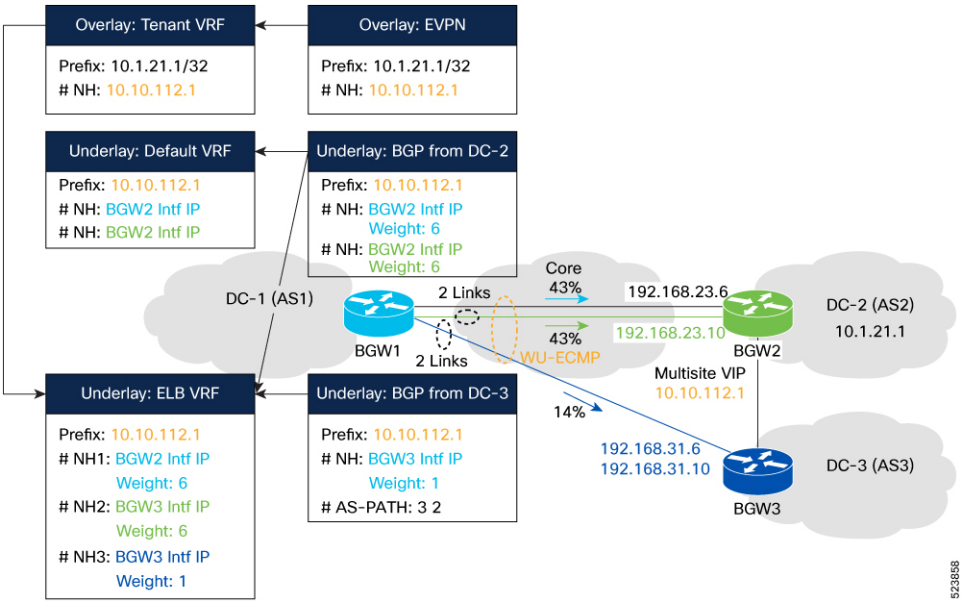
上記の出力を前の自動負荷分散の使用例で示したものと比較すると、主な違いは、明示的な負荷分散構成により、緑色のリンクの 1 つが multipath-equal-グループに含まれているため、個別に重み 3 が割り当てられることです。もう一方の緑色のリンク（multipath-equal-group に 1 つだけ残っている）にも個別の重み 3 が割り当てられますが、2 つの青色のリンクには引き続き合計重み 1 が割り当てられます。その結果、トラフィッ

クの86%が緑色のリンクを使用し（緑色のリンクごとに43%）、14%が青色のリンク（それぞれ 7%）を使用します。

緑色のパス：(6 + 6) / (6+6+1+1) = 0.86

青色のパス：(1 + 1) / (6+6+1+1) = 0.14

明示的な負荷分散ルーティングの詳細を使用した静的 wuECMP



1. **アンダーレイ ルート プログラミング**

   リモート BGW から受信したマルチパスを持つアンダーレイ ルートは、egress-loadbalance-resolution- VRF の BGP ルーティング テーブルとユニキャスト ルーティング テーブルにプログラムされます。

   以下の例では、以下のコンポーネントの出力例を示します。

   - BGP：

```
BGW1# show bgp ipv4 unicast 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.23.6 (metric 0) from 192.168.23.6 (101.2.33.1)
      Origin incomplete, MED 0, localpref 100, weight 0, explicit  weight 6

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.23.10 (metric 0) from 192.168.23.10 (101.2.33.1)
      Origin incomplete, MED 0, localpref 100, weight 0, load share weight 6

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
```

```
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
    192.168.31.6 (metric 0) from 192.168.31.6 (101.3.33.1)
      Origin incomplete, MED not set, localpref 100, weight 0, load share weight
 1

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.112.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
    192.168.31.10 (metric 0) from 192.168.31.10 (101.3.33.1)
      Origin incomplete, MED not set, localpref 100, weight 0, load share weight
 1
```

- URIB：

```
BGW1# show ip route 10.10.112.1 detail vrf egress-loadbalance-resolution-
<Truncated>
10.10.112.1/32, ubest/mbest: 4/0
    *via 192.168.23.6%default, [20/0], 1w0d, bgp-1, weight:6, external, tag 2,
uecmp ! Blue Path
<Truncated>
    *via 192.168.23.10%default, [20/0], 1w0d, bgp-1, weight:6, external, tag 2,
 uecmp ! Green Path
<Truncated>
    *via 192.168.31.6%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3,
uecmp ! Blue Path
<Truncated>
    *via 192.168.31.10%default, [20/0], 1w0d, bgp-1, weight:1, external, tag 3,
 uecmp ! Blue Path
<Truncated>
```

- FIB：

```
BGW1# show forwarding route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
*10.10.112.1/32      192.168.23.6           Ethernet1/22 >> 27 Entries
                     192.168.23.6           Ethernet1/22
                     .....
                     192.168.23.10          Ethernet1/23 >> 27 Entries
                     192.168.23.10          Ethernet1/23
                     .....
                     192.168.31.6           Ethernet1/32 >> 5 Entries
                     192.168.31.6           Ethernet1/33
                     .....
                     192.168.31.10          Ethernet1/32 >> 5 Entries
                     192.168.31.10          Ethernet1/33
                     .....
```

次に、明示的（3）、ECMP（3）、および uECMP（1）パスセットの重みに基づいて作成された FIB 転送エントリの概要を示します。

- 明示パス：27 エントリ
- ECMP パスセット：27 エントリ
- uECMP パスセット：10 エントリ

- トラフィック比率は 27:27:10 = 3:3:1 です

2. **オーバーレイ ルート プログラミング**

   オーバーレイルートは、 **egress-loadbalance-resolution-** VRF テーブルを使用して、ファブリック外部パスの VIP/PIP ネクストホップを解決します。

   以下の例では、以下のコンポーネントの出力例を示します。

   - BGP：

```
BGW1# show bgp l2vpn evpn 10.1.21.1
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, has esi_gw

             Imported to 3 destination(s)
             Imported paths list: 3001 L3-3003001 L2-2001001
  AS-Path: 2 , path sourced external to AS
    10.10.112.1 (metric 0) from 101.2.33.1 (101.2.33.1)
      Origin IGP, MED 2000, localpref 100, weight 0
      Received label 2001001 3003001
      Extcommunity: RT:1:2001001 RT:1:3003001 SOO:102.2.121.1:0
```

   - URIB：

```
BGW1# show ip route 10.1.21.1 vrf 3001
<Truncated>
10.1.21.1/32, ubest/mbest: 1/0
    *via 10.10.112.1%egress-loadbalance-resolution-, [20/2000], 00:28:03, bgp-1,
 external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x67027001 encap: VXLAN
```

   - FIB：

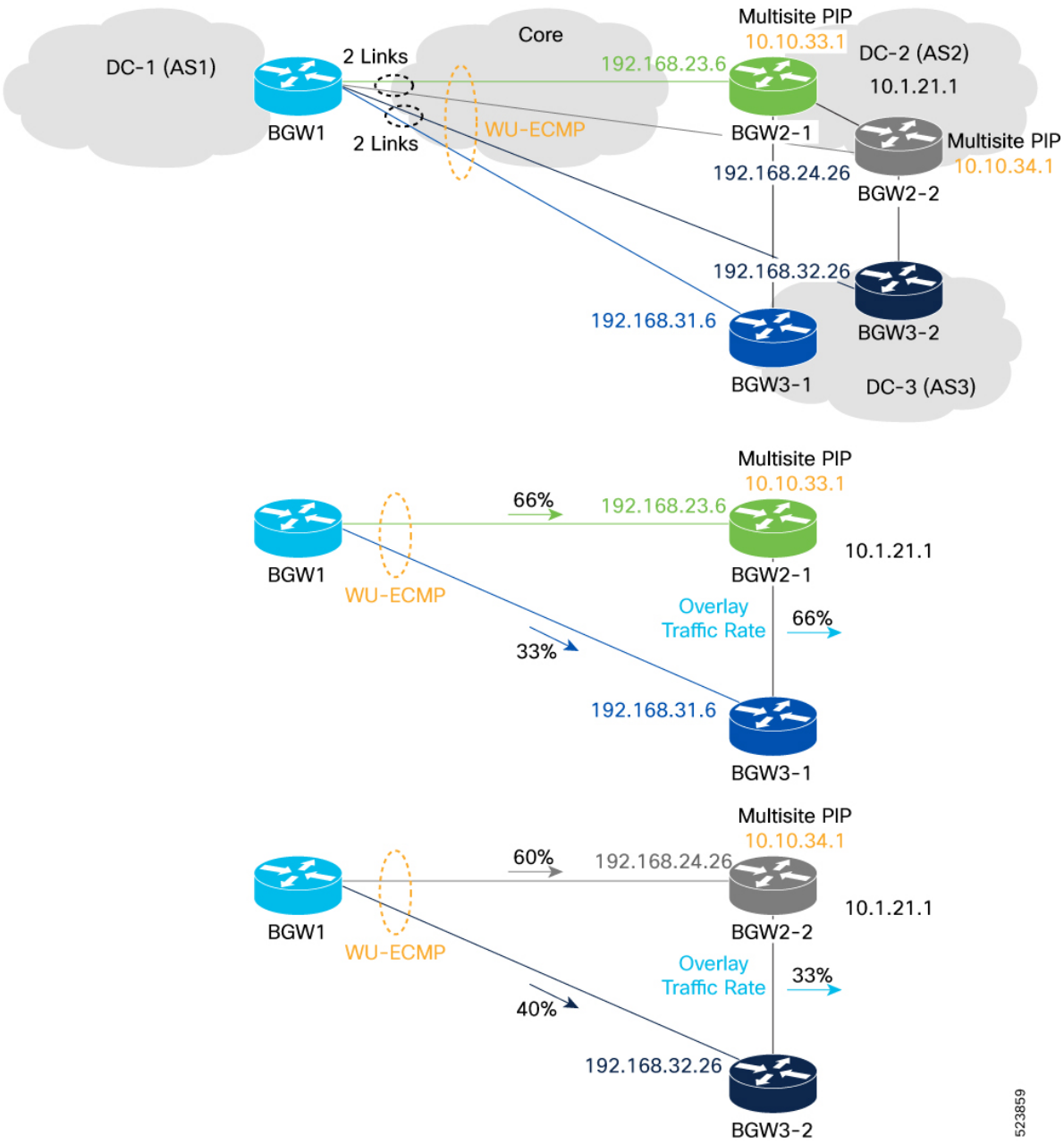
```
BGW1# show forwarding route 10.1.21.1 vrf 3001
<Truncated>
------------------+--------------------+---------------+----------------+----------------
Prefix            | Next-hop           | Interface     | Labels         |
Partial Install
------------------+--------------------+---------------+----------------+----------------
10.1.21.1/32         10.10.112.1           nve1

BGW1# show forwarding route 10.10.112.1 vrf egress-loadbalance-resolution-
<Truncated>
------------------+--------------------+---------------+----------------+----------------
Prefix            | Next-hop           | Interface     | Labels         |
Partial Install
------------------+--------------------+---------------+----------------+----------------
*10.10.112.1/32      192.168.23.6           Ethernet1/22 >> 27 Entries
                     192.168.23.6           Ethernet1/22
                     .....
                     192.168.23.10          Ethernet1/23 >> 27 Entries
                     192.168.23.10          Ethernet1/23
                     .....
                     192.168.31.6           Ethernet1/32 >> 5 Entries
                     192.168.31.6           Ethernet1/33
                     .....
                     192.168.31.10          Ethernet1/32 >> 5 Entries
                     192.168.31.10          Ethernet1/33
                     .....
```

次に、明示的（3）、ECMP（3）、および uECMP（1）パスセットの重みに基づいて作成された FIB 転送エントリの概要を示します。

- 明示パス：27 エントリ
- ECMP パスセット：27 エントリ
- uECMP パスセット：10 エントリ
- トラフィック比率は 27:27:10 = 3:3:1 です

オーバーレイとアンダーレイの動的重み（wuECMP）、およびアンダーレイの AIGP

DC-2 の BGW デバイスが 「dci-advertise-pip」コマンドで構成されることで、すべてのオーバーレイ プレフィックスが、固有のマルチサイト PIP アドレスをネクストホップとして DC-1 にアドバタイズされるシナリオを示しています。次に示す特定のトポロジのため、BGW1 は DC-2 の BGW の各マルチサイト PIP アドレスに到達するための複数のアンダーレイ パスがあり、一部はより直接的に、一部は DC-3 の BGW ノード（通常通り、それらは単に代わりにコアのルータである可能性があります）を介して間接的に到達します。前述のスタティック重みの例とは異なり、このセクションでは、AIGP メトリックの使用に基づき、よりダイナミックなアプローチを使用して、DC-2 の 2 つのマルチサイト PIP アドレス（10.10.33.1 と 10.10.34.1）に到達するために使用される異なるアンダーレイ パスに異なる重みを割り当てる方法について説明します。



アンダーレイでの AIGP を使用したダイナミック wuECMP

1. AIGP メトリックを使用してアンダーレイ ルートを発信します。
   - 関連付けられた AIGP メトリックを使用してアドバタイズする必要があるアンダーレイ ルートを指定します。この使用例では、これらのルートは DC-2 の BGW ノードのマルチサイト PIP アドレスになり、DC-1 の BGW1 にアドバタイズされるオーバーレイ プレフィックスのネクストホップを表します。

- 発信元ノード（BGW2-1/BGW2-2）では、AIGP メトリックを使用してプレフィックスを発信するために、次の 2 つのオプションを使用できます。
  - IGP コスト
  - スタティック メトリック

ここで具体的に説明されるように、アドバタイズされたプレフィックスに AIGP メトリックが含まれ、BGW PIP ループバック IP アドレスを表す場合、IGP コストを使用することで、AIGP メトリック値 0 のプレフィックスを近隣デバイスにアドバタイズすることが可能です。または、特定のメトリック値を割り当てることもできます。

次の構成例は、DC-2 の BGW ノードに適用する必要があります。

BGW2-1：

```
interface loopback1
  description #NVE_Source#
  ip address 10.10.33.1/32 tag 54321

route-map RMAP-REDIST-DIRECT permit 10
  match tag 54321
  set aigp-metric igp-cost

router bgp 2
  address-family ipv4 unicast
    redistribute direct route-map RMAP-REDIST-DIRECT
```

BGW2-2：

```
interface loopback1
  description #NVE_Source#
  ip address 10.10.34.1/32 tag 54321

route-map RMAP-REDIST-DIRECT permit 10
  match tag 54321
  set aigp-metric 4 <0 to 4294967295>

router bgp 2
  address-family ipv4 unicast
    redistribute direct route-map RMAP-REDIST-DIRECT
```

（注） **set aigp-metric** *value* または **set aigp-metric igp-cost** を使用して構成できます。ただし、両方のバリアントを同時に共存させることはできません。

2. AIGP を有効にします。
   - 発信元 BGW（BGW1/BGW3-1/BGW3-2、コアルータなど）を含むすべてのノードで、BGP ネイバーごとに BGP IPv4 AF で AIGP を有効にします。

```
router bgp 1
  template peer BGP
    address-family ipv4 unicast
      aigp
```

3. フィルタポリシーの作成

（注） ステップ 3 ～ 5 は、DC1 の BGW に適用する必要があります。

- アンダーレイ ルートを指定して、ELB wuECMP を有効にします。 アンダーレイ ルートは、DC-1 にアドバタイズされるオーバーレイ プレフィックスのネクストホップを表す、DC-2 の BGW のマルチサイト PIP アドレスです。

```
ip prefix-list site2_ms_pip1 seq 5 permit 10.10.33.1/32
ip prefix-list site2_ms_pip2 seq 5 permit 10.10.34.1/32
```

- ルートマップを作成し、照合状態で以前構成されたプレフィックスリストを適用します。

```
route-map Filter-Policy permit 10
  match ip address prefix-list site2_ms_pip1 site2_ms_pip2
```

4. ELB フィルタポリシーを有効にします（ IPv4 または IPv4 BGP アドレスファミリ下）。

```
router bgp 1
  address-family ipv4 unicast
    load-balance egress filter-policy route-map Filter-Policy
```

5. **Egress-loadbalance-resolution-** VRF テーブルのルートを確認します。次の出力に示すように、BGW1 はデフォルトで直接アンダーレイ パスを使用して DC-2 の各 PIP アドレスに到達します。

```
BGW1# show ip route 10.10.33.1 vrf egress-loadbalance-resolution-

10.10.33.1, ubest/mbest: 1/0
  *via 192.168.23.6%default, [20/4], 1d04h, bgp-1, external, tag 2, uecmp ! Green
path
BGW1# show ip route 10.10.34.1 vrf egress-loadbalance-resolution-
10.10.34.1/32, ubest/mbest: 1/0
  *via 192.168.24.26%default, [20/8], 1d04h, bgp-1, external, tag 2, uecmp ! Green
Path
```

6. マルチパス自動ポリシーを作成します。

- マルチパスを選択するときに考慮することができる AS-Path 長や AIGP メトリックなどの BGP 属性の違いを提供します。
- 「出力ロード バランシング」のために計算され、インストールされるアンダーレイ パスの最大数を指定します。

```
route-map Auto-Policy permit 10
  set as-path-length difference 1 <1 to 255>
  set aigp-metric difference 10 <1 to 4294967295>
  set maximum-paths 8 <1 to 64>
```

（注） ベスト パスと非ベスト パスの間で導出された AIGP メトリックの差が 10 未満の場合、宛先 PIP アドレスに到達するために使用されるアンダーレイリンクのマルチパスセットに非最適パスが追加されます。

7. ELB マルチパス自動ポリシーの有効化

```
router bgp 1
  address-family ipv4 unicast
    load-balance egress multipath auto-policy route-map Auto-policy
```

8. URIB **Egress-loadbalance-resolution** VRF テーブルのルートを確認します。次の出力に示すように、DC-2 の各マルチサイト PIP アドレスに到達するために 2 つの uECMP ルートがインストールされています。受信したプレフィックスについて BGW1 で計算された AIGP メトリック情報に基づいて、各パスに異なる重みが関連付けられます。

```
BGW1# show ip route 10.10.33.1 vrf egress-loadbalance-resolution-

10.10.33.1/32, ubest/mbest: 4/0
    *via 192.168.23.6%default, [20/4], 23:19:35, bgp-1, weight:2, external, tag 2,
uecmp ! Green path
    *via 192.168.31.6%default, [20/4], 23:23:59, bgp-1, weight:1, external, tag 3,
uecmp ! Blue path

BGW1# show ip route 10.10.34.1 vrf egress-loadbalance-resolution-

10.10.34.1/32, ubest/mbest: 4/0
    *via 192.168.24.26%default, [20/8], 1d00h, bgp-1, weight:3, external, tag 2,
uecmp ! Green path
    *via 192.168.32.26%default, [20/8], 1d00h, bgp-1, weight:2, external, tag 3,
uecmp ! Blue path
```

オーバーレイでの AIGP を使用した動的 wuECMP

1. 出力ロードバランシング uECMP アンダーレイ パスを使用したオーバーレイ ネクストホップ解決を有効にします。

   - このコマンドは、**Egress-loadbalance-resolution-** VRF テーブルを使用した EVPN ネクストホップの解決を有効にします。
   - アンダーレイ テーブルで出力ロード バランシング計算を有効にした後、このコマンドを有効にする必要があります。

```
router bgp 1 !BGW1
  address-family l2vpn evpn
    nexthop load-balance egress multisite
```

2. オーバーレイ wuECMP を有効にします。オーバーレイ プレフィックスがネクストホップとして PIP でアドバタイズされるようにするために、このコマンドは DC2 の BGW ノードで必要です。

```
evpn multisite border-gateway 11 !BGW2-1/BGW2-2
  dci-advertise-pip

router bgp 1 !BGW1
  address-family l2vpn evpn
    maximum-paths 8
    maximum-paths unequal-cost
  vrf 3001
    address-family ipv4 unicast
      maximum-paths 8
      maximum-paths unequal-cost
```

3. BGW1 でテナント VRF テーブルでオーバーレイ ルートを確認します。オーバーレイ プレフィックスは、両方の Multi-Site PIP アドレスをネクストホップとして使用して学習されます。上記の出力に基づいて、各 PIP アドレス宛てのトラフィックは、緑色と青色のパスを使用して（異なる重みで）不等にロードバランシングされることがわかります。さらに、次の出力に示すように、アンダーレイ AIGP メトリック情報は、ネクストホップ PIP アドレスに異なるメトリックを割り当てるためにも活用されます。それにより、どのリモート BGW ノード トラフィックをカプセル化する必要があるか決定するときに、（異なる重みで）不等ロードバランシングを適用することもできます。

```
BGW1# show ip route 10.1.21.1 vrf 3001

10.1.21.1/32, ubest/mbest: 2/0
    *via 10.10.33.1%egress-loadbalance-resolution-, [20/2000], 1d02h, bgp-1, weight:2,
 external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x66022101 encap: VXLAN
    *via 10.10.34.1%egress-loadbalance-resolution-, [20/2000], 1d02h, bgp-1, weight:1,
 external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x66022201 encap: VXLAN
```

オプションの構成

1. **bestpath aigp ignore** を有効にします。

   - 1 つのパスに AIGP メトリックがない状態で、2 つのパス間のベスト パス選択プロセスで AIGP メトリックを評価しないように、BFP を実行しているデバイスを構成する方法：

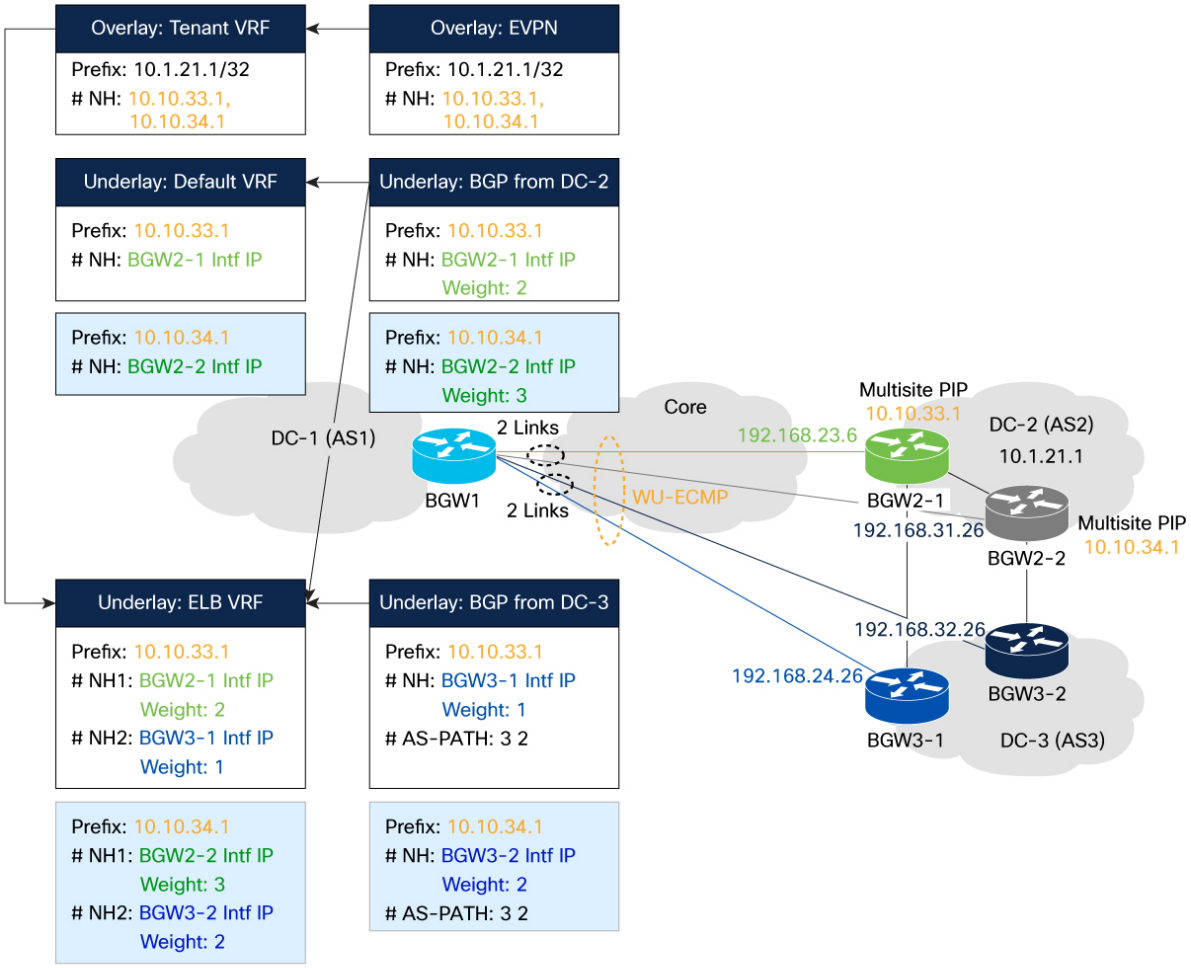
     ```
     router bgp 1
       bestpath aigp ignore
     ```

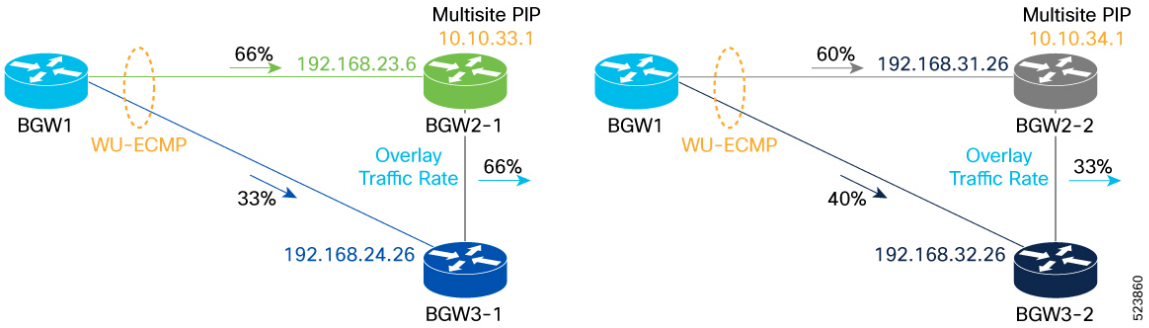
2. **reference-bandwidth**を有効にします。

   - 2 つの eBGP ピア R1 と R2 には、IGP が実行されていない直接リンクが存在する可能性があります。このようなリンクのコストを導出するには、次のコマンドを使用して参照帯域幅を構成する必要があります。

     ```
     router bgp 1
       reference-bandwidth ?
       <1-4000000>  Rate in Mbps (bandwidth) (Default)
                    *Default value is 40000
       <1-4000>     Rate in Gbps (bandwidth)
     ```

AIGP ルーティングの詳細とダイナミック wuECMP





1. **AIGP アンダーレイ ルート プログラミングを使用したダイナミック wuECMP**

   DC-2 のリモート BGW から受信したマルチパスを持つアンダーレイ ルートは、BGW1 の egress-loadbalance-resolution- VRF の BGP ルーティング テーブルとユニキャスト ルーティング テーブルにプログラムされます。

   以下の例では、以下のコンポーネントの出力例を示します。詳細については、ロードシェア重みの計算 （10 ページ）を参照してください。

- BGP：

```
BGW1# show bgp ipv4 unicast 10.10.33.1 vrf egress-loadbalance-resolution-
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
             Imported from 10.10.33.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.23.6 (metric 0) from 192.168.23.6 (10.10.33.1)
      Origin incomplete, MED 0, localpref 100, weight 0, aigp metric weight 2,
aigp 0 derived aigp = 4

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 no labeled nexthop, in rib
             Imported from 10.10.33.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.33.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
    192.168.31.6 (metric 0) from 192.168.31.6 (10.10.33.1)
     Origin incomplete, MED not set, localpref 100, weight 0, aigp metric weight
 1, aigp 4 derived aigp = 8
```

```
BGW1# show bgp ipv4 unicast 10.10.34.1 vrf egress-loadbalance-resolution-
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib
             Imported from 10.10.34.1/32 (VRF default)
  AS-Path: 2 , path sourced external to AS
    192.168.24.26 (metric 0) from 192.168.24.26 (10.10.34.1)
      Origin incomplete, MED 0, localpref 100, weight 0, aigp metric weight 3,
aigp 4 derived aigp = 8

  Path type: external, path is valid, not best reason: newer EBGP path, multipath,
 uecmp, no labeled nexthop, in rib
             Imported from 10.10.34.1/32 (VRF default)
  AS-Path: 3 2 , path sourced external to AS
    192.168.32.26 (metric 0) from 192.168.32.26 (10.10.34.1)
     Origin incomplete, MED not set, localpref 100, weight 0, aigp metric weight
 2, aigp 8 derived aigp = 12
```

- URIB：

```
BGW1# show ip route 10.10.33.1 detail vrf egress-loadbalance-resolution-
<Truncated>
10.10.33.1/32, ubest/mbest: 4/0
    *via 192.168.23.6%default, [20/4], 23:19:35, bgp-1, weight:2, external, tag
 2, uecmp ! Green path
<Truncated>
    *via 192.168.31.6%default, [20/4], 23:23:59, bgp-1, weight:1, external, tag
 3, uecmp ! Blue path
<Truncated>
```

```
BGW1# show ip route 10.10.34.1 detail vrf egress-loadbalance-resolution-
<Truncated>
10.10.34.1/32, ubest/mbest: 4/0
    *via 192.168.24.26%default, [20/8], 1d00h, bgp-1, weight:3, external, tag 2,
 uecmp ! Green path
<Truncated>
    *via 192.168.32.26%default, [20/8], 1d00h, bgp-1, weight:2, external, tag 3,
```

```
 uecmp ! Blue path
<Truncated>
```

- FIB：

```
BGW1# show forwarding route 10.10.33.1 vrf egress-loadbalance-resolution-
<Truncated>
------------------+--------------------+---------------+----------------+----------------
Prefix            | Next-hop           | Interface     | Labels         |
Partial Install
------------------+--------------------+---------------+----------------+----------------
*10.10.33.1/32       192.168.23.6           Ethernet1/22 ! 21 Entries
                     192.168.23.6           Ethernet1/22
                     .....
                     192.168.31.6           Ethernet1/32 ! 11 Entries
                     192.168.31.6           Ethernet1/33
                     .....
```

次に、ECMP（2）およびuECMP（1）パスセットの重みに基づいて作成されたFIB転送エントリの概要を示します。

- ECMP パスセット：21 エントリ
- uECMP パスセット：11 エントリ
- トラフィック比率は 21 : 11 = 2 : 1

```
BGW1# show forwarding route 10.10.34.1 vrf egress-loadbalance-resolution-
<Truncated>
------------------+--------------------+---------------+----------------+----------------
Prefix            | Next-hop           | Interface     | Labels         |
Partial Install
------------------+--------------------+---------------+----------------+----------------
*10.10.34.1/32       192.168.24.26          Ethernet1/27 ! 19 Entries
                     192.168.24.26          Ethernet1/27
                     .....
                     192.168.32.26          Ethernet1/37 ! 13 Entries
                     192.168.32.26          Ethernet1/37
                     .....
```

次に、ECMP（3）およびuECMP（2）パスセットの重みに基づいて作成されたFIB転送エントリの概要を示します。

- ECMP パスセット：19 エントリ
- uECMP パスセット：13 エントリ
- トラフィック比率は 19 : 13 = 3 : 2 です。

2. **AIGP オーバーレイ ルート プログラミングを使用したダイナミック wuECMP**

   オーバーレイ ルートは、BGW1 の **egress-loadbalance-resolution-** VRF テーブルを使用して、DC-2から受信したオーバーレイ プレフィックスのPIPネクストホップを解決します。

   以下の例では、以下のコンポーネントの出力例を示します。詳細については、アンダーレイのダイナミックの導出 （9 ページ）を参照してください。

   - BGP：

```
BGW1# show bgp l2vpn evpn 10.1.21.1
<Truncated>
  Advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, has esi_gw

             Imported to 3 destination(s)
             Imported paths list: 3001 L3-3003001 L2-2001001
  AS-Path: 2 , path sourced external to AS
    10.10.33.1 (metric 4) from 10.10.33.1 (10.10.33.1)
<Truncated>
  Path type: external, path is valid, not best reason: NH metric, multipath, no
 labeled nexthop, has esi_gw
             Imported to 3 destination(s)
             Imported paths list: 3001 L3-3003001 L2-2001001
  AS-Path: 2 , path sourced external to AS
    10.10.34.1 (metric 8) from 10.10.34.1 (10.10.34.1)

BGW1# show bgp ipv4 unicast 10.1.21.1 vrf 3001
<Truncated>
  Advertised path-id 1, VPN AF advertised path-id 1
  Path type: external, path is valid, is best path, no labeled nexthop, in rib,
 has esi_gw
             Imported from
21:2001001:[2]:[0]:[0]:[48]:[0010.0100.2101]:[32]:[10.1.21.1]/272
  AS-Path: 2 , path sourced external to AS
    10.10.33.1 (metric 4) from 10.10.33.1 (10.10.33.1)
      Origin IGP, MED 2000, localpref 100, weight 0, igp metric weight 2
<Truncated>
  Path type: external, path is valid, not best reason: NH metric, multipath, no
 labeled nexthop, in rib, has esi_gw
             Imported from
21:2001001:[2]:[0]:[0]:[48]:[0010.0100.2101]:[32]:[10.1.21.1]/272
  AS-Path: 2 , path sourced external to AS
    10.10.34.1 (metric 8) from 10.10.34.1 (10.10.34.1)
      Origin IGP, MED 2000, localpref 100, weight 0, igp metric weight 1
```

- URIB：

```
BGW1# show ip route 10.1.21.1 detail vrf 3001
<Truncated>
10.1.21.1/32, ubest/mbest: 2/0
    *via 10.10.33.1%egress-loadbalance-resolution-, [20/2000], 1d02h, bgp-1,
weight:2, external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x66022101 encap: VXLAN
<Truncated>
    *via 10.10.34.1%egress-loadbalance-resolution-, [20/2000], 1d02h, bgp-1,
weight:1, external,
    tag 2, eLB, segid: 3003001 tunnelid: 0x66022201 encap: VXLAN
```

- FIB：

```
BGW1# show forwarding route 10.1.21.1 vrf 3001
<Truncated>
-----------------+--------------------+---------------+----------------+---------------
Prefix           | Next-hop           | Interface     | Labels         |
Partial Install
-----------------+--------------------+---------------+----------------+---------------
10.1.21.1/32        10.10.33.1          nve1 ! 21 Entries
                    10.10.33.1          nve1
                    .....
                    10.10.34.1          nve1 ! 11 Entries
                    10.10.34.1          nve1
                    .....
```

- L2RIB：重み付きの Mac ルート（EVPN タイプ 2）：

```
BGW1# show l2route evpn mac evi 1001 detail | be "0010.0100.2101"
1001       0010.0100.2101 BGP    SplRcv            0         10.10.33.1 (Label:
 2001001)
                                                            10.10.34.1 (Label:
 2001001)
            Route Resolution Type: ESI
            Forwarding State: Resolved (PL)
            Resultant PL: 10.10.33.1(Wt: 2), 10.10.34.1(Wt: 1)
            Sent To: L2FM

BGW1# show l2route evpn path-list all detail
<Truncated>
Topology ID  Prod   ESI                       ECMP Label Flags  Client Ctx  MACs
       NFN Bitmap
------------ ------ ------------------------- ---------- ------ -----------
---------- ----------
<Truncated>
1001         UFDM   0300.0000.0000.1500.0309  2          A      1493174825  0
        4
                     CP Next-Hops: 10.10.33.1, 10.10.34.1
                     Gbl EAD Next-Hops:
                     Res Next-Hops: 10.10.33.1(Wt: 2), 10.10.34.1(Wt: 1)
                     Bkp Next-Hops:
```

- 重み付き L2FM - Mac ルート（EVPN タイプ 2）：

```
BGW1# show mac address-table vlan 1001 address 0010.0100.2101
<Truncated>
 VLAN     MAC Address      Type      age     Secure NTFY Ports
---------+-----------------+--------+---------+------+----+------------------
C 1001     0010.0100.2101   dynamic  NA         F      F    nve1(10.10.33.1[Wt:
 2] 10.10.34.1[Wt: 1])
```

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。