



セグメント ルーティングでの出力ピア エンジニアリング

- [BGP プレフィックス SID \(1 ページ\)](#)
- [隣接 SID \(1 ページ\)](#)
- [セグメント ルーティングのための高可用性 \(2 ページ\)](#)
- [セグメント ルーティングを使用した BGP 出力ピア エンジニアリングの概要 \(2 ページ\)](#)
- [BGP 出力ピア エンジニアリングのガイドラインと制限事項 \(4 ページ\)](#)
- [BGP を使用したネイバー出力ピア エンジニアリングの設定 \(4 ページ\)](#)
- [出力ピア エンジニアリングの設定例 \(6 ページ\)](#)
- [BGP リンク ステート アドレス ファミリの設定 \(8 ページ\)](#)
- [BGP プレフィックス SID の展開例 \(9 ページ\)](#)

BGP プレフィックス SID

セグメント ルーティングをサポートするためには、BGP が BGP プレフィックスのセグメント ID (SID) をアドバタイズできなければなりません。BGP プレフィックス SID は常にセグメント ルーティング BGP ドメイン内でグローバルであり、命令を識別し、BGP によって計算された ECMP 対応のベストパスを介して、パケットを関連するプレフィックスに転送します。BGP プレフィックス SID は、BGP プレフィックス セグメントを識別します。

隣接 SID

隣接関係セグメント識別子 (SID) は、特定のインターフェイスとそのインターフェイスからの次のホップを指す、ローカル ラベルです。隣接関係 SID を有効にするために必要な特定の設定はありません。アドレス ファミリの BGP を介してセグメント ルーティングが有効になると、BGP が実行されるすべてのインターフェイスに対して、アドレスファミリーがそのインターフェイスのすべてのネイバーに対して隣接 SID を自動的に割り当てます。

セグメントルーティングのための高可用性

インサービス ソフトウェア アップグレード (ISSU) は、BGP グレースフル リスタートで最低限サポートされます。すべての状態（セグメントルーティング状態を含む）は、BGP ルータのピアから再学習する必要があります。グレースフルリスタート期間中、以前に学習したルートとラベルの状態は保持されます。

セグメントルーティングを使用した BGP 出力ピア エンジニアリングの概要

Cisco Nexus 9000 シリーズ スイッチは、多くの場合、大規模データセンター（MSDC）に導入されます。このような環境では、セグメントルーティング（SR）で BGP 出力ピア エンジニアリング（EPE）をサポートすることが要件となります。

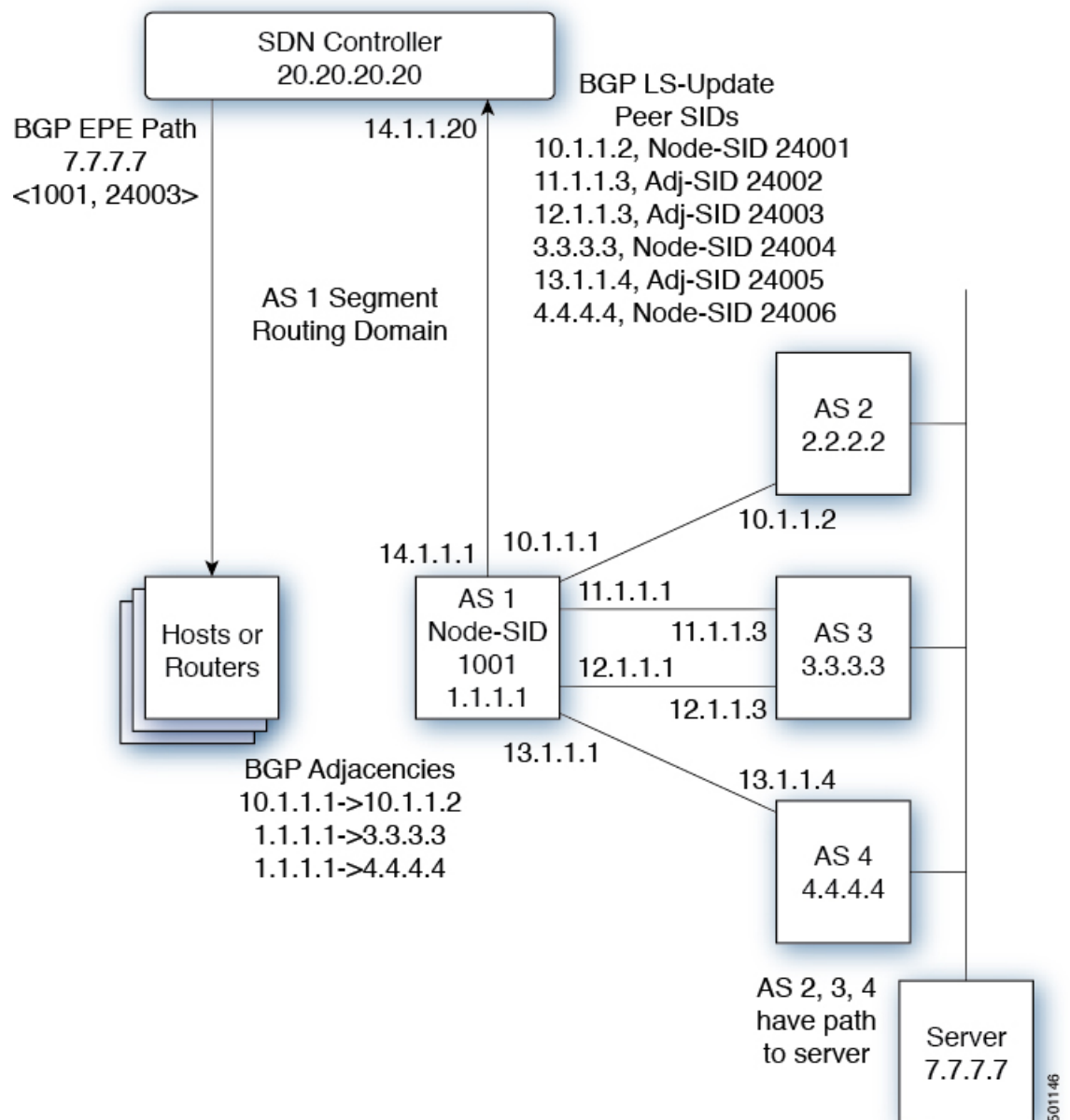
セグメントルーティング（SR）はソースルーティングを利用します。ノードは、制御された一連の命令（セグメント）によってパケットを操作するために、パケットの前に SR ヘッダーを付加します。セグメントは、トポロジまたはサービススペースの命令を表すことができます。SR では、SR ドメインの入力ノードでのみフローごとの状態を維持しながら、トポロジパスまたはサービスチェーンを介してフローを操作できます。この機能の場合、セグメントルーティングアーキテクチャは、MPLS データプレーンに直接適用されます。

セグメントルーティングをサポートするためには、BGP が BGP プレフィックスのセグメント ID（SID）をアドバタイズできなければなりません。BGP プレフィックスは常に SR または BGP ドメイン内でグローバルであり、命令を識別し、BGP によって計算された ECMP 対応のベストパスを介して、パケットを関連するプレフィックスに転送します。BGP プレフィックスは、BGP プレフィックス セグメントの識別子です。

SR ベースの出力ピア エンジニアリング（EPE）ソリューションにより、集中型（SDN）コントローラは、ドメイン内の入力境界ルータまたはホストで任意の出力ピアポリシーをプログラムできます。

次の例では、3 つのルータすべてが iBGP を実行し、NRLI を相互にアドバタイズします。また、ルータはループバックをネクストホップとしてアドバタイズし、再帰的に解決します。これにより、図に示すように、ルータ間に ECMP が提供されます。

図 1: 出力ピア エンジニアリングの例



SDN コントローラは、そのピアおよび隣接のそれぞれについて、出力ルータ 1.1.1.1 からのセグメント ID を受信します。次に、出口ポイントをコントローラのルーティング ドメイン内の他のルータおよびホストにインテリジェントにアダプタイズできます。図に示すように、BGP ネットワーク層到達可能性情報 (NLRI) には、ルータ 1.1.1.1 へのノード SID と、7.7.7.7 へのトラフィックがリンク 12.1.1.1->12.1.1.3 を介して出力されることを示すピア隣接 SID 24003 の両方が含まれています。

BGP 出力ピアエンジニアリングのガイドラインと制限事項

BGP 出力ピアエンジニアリングには、次のガイドラインと制限事項があります。

- BGP 出力ピアエンジニアリングは、IPv4 BGP ピアでのみサポートされています。IPv6 BGP ピアはサポートされていません。
- BGP 出力ピアエンジニアリングは、デフォルトの VPN ルーティングおよび転送（VRF）インスタンスでのみサポートされます。
- 出力ピアエンジニアリング（EPE）ピアセットには、任意の数の EPG ピアを追加できます。ただし、インストールされている復元力のある CE ごとの FEC は 32 ピアに制限されています。
- 特定の BGP ネイバーは、単一のピアセットのメンバーにしかありません。ピアセットが構成されています。複数のピアセットはサポートされていません。オプションのピアセット名を指定して、ネイバーをピアセットに追加できます。対応する RPCFEC は、ピアセット内のすべてのピア間でトラフィックを負荷分散します。ピアセット名は、最長 63 文字の文字列です（64 NULL で終了）。この長さは、NX-OS ポリシー名の長さと一致します。ピアは、単一のピアセットのメンバーにしかありません。
- 特定のピアの隣接関係は、異なるピアセットに個別に割り当てることはできません。
- Cisco NX-OS リリース 9.3(3) 以降、BGP 出力ピアエンジニアリングは Cisco Nexus 9300-GX プラットフォーム スイッチでサポートされます。

BGP を使用したネイバー出力ピアエンジニアリングの設定

RFC 7752 および draft-ietf-idr-bgppls-segment-routing-epe の導入により、出力園児に名リングを設定できます。この機能は、外部 BGP ネイバーに対してのみ有効であり、デフォルトでは設定されていません。出力エンジニアリングでは、RFC 7752 エンコーディングを使用します。

始める前に

- BGP を有効にする必要があります。
- リリース 7.0(3)I3(1) またはリリース 7.0(3)I4(1) からアップグレードした後、Cisco Nexus 9000 シリーズ スイッチで出力ピアエンジニアリング（EPE）を設定する前に、次のコマンドを使用して、TCAM リージョンを設定します。
 1. switch# **hardware access-list tcam region vpc-convergence 0**
 2. switch# **hardware access-list tcam region racl 0**

3. switch# hardware access-list tcam region mpls 256 double-wide

- 設定を保存して、スイッチをリロードします。

詳細については、*Cisco Nexus 9000 Series NX-OS Security Configuration Guide* の「Using Templates to Configure ACL TCAM Region Sizes」および「Configuring ACL TCAM Region Sizes」のセクションを参照してください。

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# configure 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 <bgp autonomous number>	自律ルータ BGP 番号を指定します。
ステップ 3	neighbor <IP address>	ネイバーの IP アドレスを設定します。
ステップ 4	[no default] egress-engineering [peer-set peer-set-name] 例 : switch(config)# router bgp 1 switch(config-router)# neighbor 4.4.4.4 switch(config-router)# egress-engineering peer-set NewPeer	ピアノード SID がネイバーに割り当てられ、BGP リンク状態 (BGP-LS) アドレス ファミリー リンク NLRI のインスタンスでアドバタイズされるかどうかを指定します。ネイバーがマルチホップ ネイバーである場合、BGP-LS リンク NLRI インスタンスもネイバーへの等コスト マルチパス (ECMP) パスごとにアドバタイズされます。これには、一意の Peer-Adj-SID が含まれます。 オプションで、ネイバーをピアセットに追加できます。ピアセット SID は、ピアノード SID と同じインスタンスの BGP-LS リンク NLRI でもアドバタイズされます。BGP リンクステート NLRI は、リンクステートアドレス ファミリーが設定されているすべてのネイバーにアドバタイズされます。 EPE の詳細については、RFC 7752 および draft-ietf-idr-bgpls-segment-routing-epe-05 を参照してください。

出力ピアエンジニアリングの設定例

BGP スピーカー 1.1.1.1 の出力ピア エンジニアリングのサンプル設定を参照してください。ネイバー 20.20.20.20 は SDN コントローラであることに注意してください。

```
hostname epe-as-1
install feature-set mpls
feature-set mpls

feature telnet
feature bash-shell
feature scp-server
feature bgp
feature mpls segment-routing

segment-routing mpls
vlan 1

vrf context management
  ip route 0.0.0.0/0 10.30.97.1
  ip route 0.0.0.0/0 10.30.108.1

interface Ethernet1/1
  no switchport
  ip address 10.1.1.1/24
  no shutdown

interface Ethernet1/2
  no switchport
  ip address 11.1.1.1/24
  no shutdown

interface Ethernet1/3
  no switchport
  ip address 12.1.1.1/24
  no shutdown

interface Ethernet1/4
  no switchport
  ip address 13.1.1.1/24
  no shutdown

interface Ethernet1/5
  no switchport
  ip address 14.1.1.1/24
  no shutdown

interface mgmt0
  ip address dhcp
  vrf member management

interface loopback1
  ip address 1.1.1.1/32
  line console

line vty
  ip route 2.2.2.2/32 10.1.1.2
  ip route 3.3.3.3/32 11.1.1.3
  ip route 3.3.3.3/32 12.1.1.3
  ip route 4.4.4.4/32 13.1.1.4
```

```

ip route 20.20.20.20/32 14.1.1.20

router bgp 1
  address-family ipv4 unicast
  address-family link-state
  neighbor 10.1.1.2
    remote-as 2
    address-family ipv4
    egress-engineering
  neighbor 3.3.3.3
    remote-as 3
    address-family ipv4
    update-source loopback1
    ebgp-multihop 2
    egress-engineering
  neighbor 4.4.4.4
    remote-as 4
    address-family ipv4
    update-source loopback1
    ebgp-multihop 2
    egress-engineering
  neighbor 20.20.20.20
    remote-as 1
    address-family link-state
    update-source loopback1
    ebgp-multihop 2
  neighbor 124.11.50.5
    bfs
    remote-as 6
    update-source port-channel50.11
    egress-engineering peer-set pset2 <<<<<<
    address-family ipv4 unicast
  neighbor 124.11.101.2
    bfd
    remote-as 6
    update-source Vlan2401
    egress-engineering
    address-family ipv4 unicast

```

次に、**show bgp internal epe** コマンドの出力例を示します。

```

switch# show bgp internal epe
BGP Egress Peer Engineering (EPE) Information:
Link-State Server: Inactive
Link-State Client: Active
Configured EPE Peers: 26
Active EPE Peers: 3
EPE SID State:
RPC SID Peer or Set Assigned
ID Type Set Name ID Label Adj-Info, iod
1 Node 124.1.50.5 1 1600
2 Set pset1 2 1601
3 Node 6.6.6.6 3 1602
4 Node 124.11.50.5 4 1603
5 Set pset2 5 1604
6 Adj 6.6.6.6 6 1605 124.11.50.4->124.11.50.5/0x1600b031, 80
7 Adj 6.6.6.6 7 1606 124.1.50.4->124.1.50.5/0x16000031, 78
EPE Peer-Sets:
IPv4 Peer-Set: pset1, RPC-Set 2, Count 7, SID 1601
Peers: 124.11.116.2 124.11.111.2 124.11.106.2 124.11.101.2
124.11.49.5 124.1.50.5 124.1.49.5
IPv4 Peer-Set: pset2, RPC-Set 5, Count 5, SID 1604
Peers: 124.11.117.2 124.11.112.2 124.11.107.2 124.11.102.2
124.11.50.5

```

```
IPv4 Peer-Set: pset3, RPC-Set 0, Count 4, SID unspecified
Peers: 124.11.118.2 124.11.113.2 124.11.108.2 124.11.103.2
IPv4 Peer-Set: pset4, RPC-Set 0, Count 4, SID unspecified
Peers: 124.11.119.2 124.11.114.2 124.11.109.2 124.11.104.2
IPv4 Peer-Set: pset5, RPC-Set 0, Count 4, SID unspecified
Peers: 124.11.120.2 124.11.115.2 124.11.110.2 124.11.105.2
switch#
```

BGP リンク ステート アドレス ファミリの設定

対応する SID をアドバタイズするコントローラを持つネイバーセッションに対し、BGP リンク ステート アドレス ファミリを設定することができます。この機能は、グローバル コンフィギュレーション モードおよびネイバー アドレス ファミリ コンフィギュレーション モードで設定できます。

始める前に

BGP を有効にする必要があります。

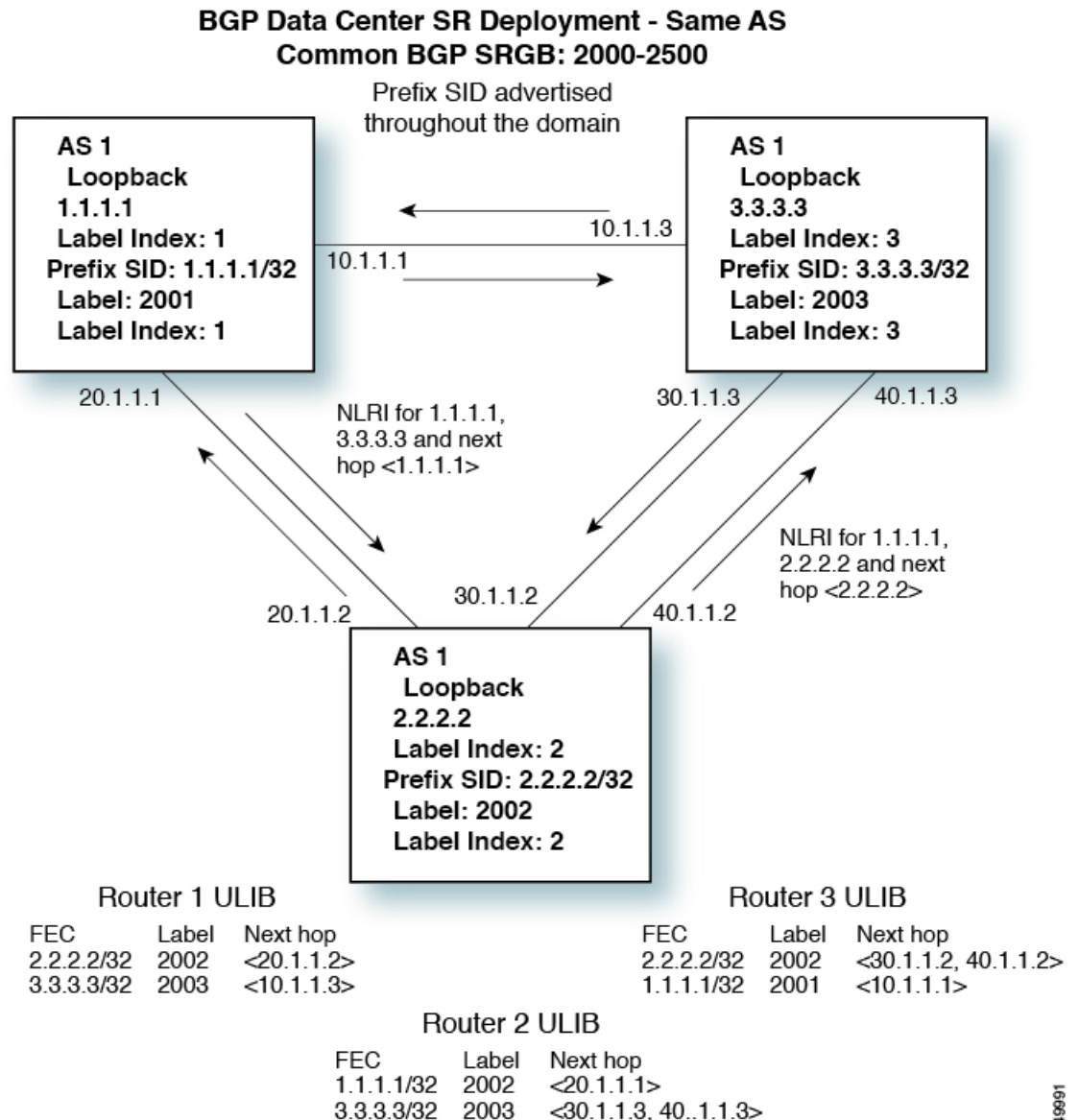
手順

	コマンドまたはアクション	目的
ステップ 1	configure terminal 例 : switch# configure 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 <bgp autonomous number>	自律ルータ BGP 番号を指定します。
ステップ 3	[no] address-family link-state 例 : switch(config)# router bgp 64497 switch (config-router af)# address-family link-state	アドレスファミリ インターフェイス コンフィギュレーション モードを開始します。 (注) このコマンドは、ネイバー アドレス ファミリ コンフィギュレーション モードでも設定できます。
ステップ 4	neighbor <IP address>	ネイバーの IP アドレスを設定します。
ステップ 5	[no] address-family link-state 例 : switch(config)#router bgp 1 switch(config-router)#address-family link-state switch(config-router)#neighbor 20.20.20.20 switch(config-router)#address-family link-state	アドレスファミリ インターフェイス コンフィギュレーション モードを開始します。 (注) このコマンドは、ネイバー アドレス ファミリ コンフィギュレーション モードでも設定できます。

BGP プレフィックス SID の展開例

以下の簡単な例では、3 つのルーターすべてが iBGP を実行し、ネットワーク層到達可能性情報 (NLRI) を互いにアドバタイズしています。また、ルーターは、ルーター 2.2.2.2 と 3.3.3.3 の間に ECMP を提供するネクスト ホップとして、ループバック インターフェイスをアドバタイズしています。

図 2: BGP プレフィックス SID の簡単な例



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。