

BGP で最適パスを選択するアルゴリズム

目次

[概要](#)

[ルータでパスが無視される理由](#)

[最適パス アルゴリズムの仕組み](#)

[例：BGP 最良パスの選択](#)

[パスを選択するプロセスのカスタマイズ](#)

[BGP マルチパス](#)

[関連情報](#)

概要

Border Gateway Protocol (BGP; ボーダーゲートウェイ プロトコル) ルータは、通常は同じ宛先へのパスを複数受信します。BGP の最適パス アルゴリズムは、IP ルーティング テーブルに格納し、トラフィック転送に使用する最適なパスを決めるものです。

ルータでパスが無視される理由

特定のプレフィクスについてルータで受信されるすべてのパスがリストに並んでいると仮定します。リストは [提示 IP bgplonger](#) プレフィクス コマンドの出力に類似したです。この場合、パスの中には最適パスの候補と見なされないものがあります。そのようなパスには通常、`show ip bgp longer-prefixes` コマンドの出力中に有効なフラグがありません。ルータでは、次の状況にあるパスは無視されます。

- `show ip bgp longer-prefixes` 出力でとして表示されるパス。BGP 同期がイネーブルになっている場合、内部 BGP (iBGP) パスが有効なパスと見なされるためには、IP ルーティング テーブル内にプレフィクスの一致するパスが必要です。Cisco IOS(R) ソフトウェアでは、デフォルトで BGP 同期がイネーブルになっています。一致するルートを開 Shortest Path First (OSPF) 隣接ルータから学習した場合、その OSPF ルータの ID と、iBGP 隣接ルータの BGP ルータ ID とが必ず一致します。ほとんどのユーザは、[no synchronization BGP サブコマンド](#)を使用して同期を無効にすることを好みます。注: Cisco IOS 12.2(8)T
- NEXT_HOP にアクセスできないパス。パスに関連付けられている NEXT_HOP への Interior Gateway Protocol (IGP) ルートが存在することを確認します。
- ローカル自律 システム (AS) が AS_PATH に現われる場合 external BGP (eBGP) ネイバーからのパス。このようなパスはルータへの入力時に拒否され、BGP Routing Information Base (RIB; ルーティング情報ベース) にも格納されません。同じことは、アクセス、プレフィクス、AS_PATH、またはコミュニティ リストを介して実装されたルーティング ポリシーによって拒否されるパスについても当てはまります。ただし、ネイバーに対して [neighbor soft-reconfiguration inbound](#) を設定している場合は除きます。
- AS_SEQUENCE で最初の AS 数としてネイバー現在でイネーブルになった [bgp enforce-first-as](#) およびアップデート含まれなければ。この場合、ルータによって通知が送信され、セッションが閉じられます。
- `show ip bgp longer-prefixes` の出力で (received-only) とマークされているパスこれらのパスはポリシーに基づいて拒否されています。ただし、パスを送信しているネイバーに対して `soft-`

reconfiguration inbound が設定されているため、これらのパスはルータによって格納されています。

最適パス アルゴリズムの仕組み

BGP は、最初の有効なパスを現在の最適なパスとして割り当てます。そして、有効なパスのリストの最後に達するまで、BGP によって最適なパスとリスト内の次のパスとが比較されます。このリストには最適なパスの決定に使用されるルールが規定されています。

1. 最も高い WEIGHT を持つパスが優先されます。注: [WEIGHT](#) は シスコ独自のパラメータです。設定されているルータに対してローカルに割り当てられます。
2. 最も高い [LOCAL_PREF](#) を持つパスが優先されます。注: LOCAL_PREF のないパスは [bgp デフォルト](#) ローカル preference command の値が設定が、または 100 という値を持つためにデフォルトであると考慮されます。
3. `network` または `aggregate` BGP サブコマンドによって、あるいは IGP からの再配布を通じて、ローカルで発信されたパスが優先されます。[network](#) コマンドや `redistribute` コマンドによるローカル パスの方が、[aggregate-address](#) コマンドによるローカル集約よりも優先されます。注: この項目を理解しておいて下さい:

- AIGP が設定され、[bgp bestpath aigp ignore コマンド](#)が設定されなければ、決定プロセスは AIGP メトリックを考慮します。更に詳しい情報については [BGP については設定を AIGP Metric 属性](#)参照して下さい。

4. 最短の AS_PATH を持つパスが優先されます。注: 次の項目にも注意してください。

-このステップは [bgp bestpath as-path ignore](#) コマンドを設定する場合スキップされます。

- AS_SET は 1 としていくつかの AS がセットにあっても、数えます。

- AS_CONFED_SEQUENCE および AS_CONFED_SET は AS_PATH 長さに含まれていません。

5. 最小のオリジン タイプを持つパスが優先されます。注: IGP は Exterior Gateway Protocol (EGP) よりも小さく、EGP は INCOMPLETE よりも小さくなります。
6. 最小の [Multi-Exit Discriminator \(MED \)](#) を持つパスが優先されます。注: 次の項目にも注意してください。

-この比較は最初の (近隣) AS が 2 つのパスに同じである場合その時だけ発生します。コンフェデレーション サブ AS はすべて無視されます。

つまり、AS_SEQUENCE 内の最初の AS が複数のパスの間で同じ場合にのみ、MED が比較されます。それより前にある AS_CONFED_SEQUENCE はすべて無視されます。

- [bgp always-compare-med](#) が有効になれば、MED はすべてのパスのために比較されます。このオプションを AS 全体にわたってディセーブルにする必要があります。そうしないと、ルーティング ループが発生する可能性があります。

- [bgp bestpath med-confed](#) が有効になれば、MED は AS_CONFED_SEQUENCE からだけ成っているすべてのパスのために比較されます。

これらのパスは、ローカル コンフェデレーション内で発信されています。

- 4,294,967,295 の MED のネイバーから届くパスの MED は BGP テーブルに挿入の前に変更されます。MED は 4,294,967,294 に変更されます。

- 4,294,967,295 の MED のネイバーから届くパスの MED は有効と見なされ、効果の BGP テーブルに Cisco バグ ID [CSCef34800](#) のために固定されるコードに挿入されます。

- [bgp bestpath med missing-as-worst](#) を有効にしなかったら、MED 無しで受け取ったパス 0 の MED は割り当てられます。

[bgp bestpath med missing-as-worst](#) を有効にしている場合、パスには 4,294,967,294 の MED が割り当てられます。

Cisco Bug ID [CSCef34800](#) で修正されたコードにより、[bgp bestpath med missing-as-worst](#) を有効にしている場合、パスには 4,294,967,295 の MED が割り当てられます。

- [bgp deterministic-med コマンド](#) はまたこのステップに影響を及ぼすことができます。

実例については、『[BGP ルータでの最適パスの選択のための Multi-Exit Discriminator 使用方法](#)』を参照してください。

7. iBGP パスよりも eBGP パスの方が優先されます。最適パスが選択される場合は、ステップ 9 に移動してください (マルチパス)。注: AS_CONFED_SEQUENCE と AS_CONFED_SET が含まれているパスはコンフェデレーションに対してローカルなものです。したがって、これらのパスは内部パスとして扱われます。Confederation External と Confederation Internal との間には違いはありません。
8. BGP ネクストホップへの最小の IGP メトリックを持つパスが優先されます。最適パスがすでに選択されていても、続けてください。
9. マルチパスが [BGP マルチパス](#) 用にルーティング テーブルでインストールされる必要があるかどうか判断します。最適パスがまだ選択されていない場合、続けてください。
10. 両方のパスが外部のときは、先に受信したパス (最も古いパス) が優先されます。この手順によってルートフラップが最小限に抑えられます。これは、たとえ次の決定条件 (手順 11、12、および 13) に基づいて新しい方のパスが優先ルートになった場合でも、新しい方のパスによって古い方のパスが置き換えられないためです。これらの項目のいずれかが該当する場合、この手順をスキップします。[bgp best path compare-routerid コマンドをイネーブルにしている場合](#)。注: このコマンドは、Cisco IOS ソフトウェア リリース 12.0.11S、12.0.11SC、12.0.11S3、12.1.3、12.1.3AA、12.1.3.T、および 12.1.3.E で導入されています。ルータ ID が複数のパスで同じ場合 (ルートを同じルータから受信したため)。現時点で最適なパスがない場合。たとえば、パスを供給している隣接ルータがダウンした場合などには、現時点での最適なパスが失われる可能性があります。
11. 最小のルータ ID を持つ BGP ルータから送られたルートが優先されます。ルータ ID はルータ上で最大の IP アドレスで、ループバック アドレスに割り当てられた優先度が備わっています。また、[bgp router-id コマンド](#) を使用して、ルータ ID を手動で設定することもできます。注: Route ReflectorRR; ID ID
12. 発信元 ID またはルータ ID が複数のパスで同じ場合は、最小のクラスタ リスト長を持つパスが優先されます。これが発生するのは、BGP RR 環境でだけです。この環境では、クライアントは RR または他のクラスタのクライアントとピア関係を確立できます。このシナリオでは、クライアントは RR 固有の BGP アトリビュートを認識している必要があります。
13. 最小の隣接ルータ アドレスから送られたパスが優先されます。このアドレスは BGP ネイバー設定で使用されている IP アドレスです。このアドレスはローカル ルータとの TCP 接続に使用されるリモートピアに対応します。

例： BGP 最良パスの選択

この例では、ネットワーク 10.30.116.0/23 に対して 9 つのパスが使用可能です。 `show ip bgp network` コマンドは、特定のネットワークの BGP ルーティング テーブルのエントリを表示します。

```
Router R1#show ip bgp vpnv4 rd 1100:1001 10.30.116.0/23
BGP routing table entry for 1100:1001:10.30.116.0/23, version 26765275
Paths: (9 available, best #6, no table)
  Advertised to update-groups:
    1          2          3
(65001 64955 65003) 65089, (Received from a RR-client)
  172.16.254.226 (metric 20645) from 172.16.224.236 (172.16.224.236)
    Origin IGP, metric 0, localpref 100, valid, confed-internal
    Extended Community: RT:1100:1001
    mpls labels in/out nolabel/362
(65008 64955 65003) 65089
  172.16.254.226 (metric 20645) from 10.131.123.71 (10.131.123.71)
    Origin IGP, metric 0, localpref 100, valid, confed-external
    Extended Community: RT:1100:1001
    mpls labels in/out nolabel/362
(65001 64955 65003) 65089
  172.16.254.226 (metric 20645) from 172.16.216.253 (172.16.216.253)
    Origin IGP, metric 0, localpref 100, valid, confed-external
    Extended Community: RT:1100:1001
    mpls labels in/out nolabel/362
(65001 64955 65003) 65089
  172.16.254.226 (metric 20645) from 172.16.216.252 (172.16.216.252)
    Origin IGP, metric 0, localpref 100, valid, confed-external
    Extended Community: RT:1100:1001
    mpls labels in/out nolabel/362
(64955 65003) 65089
  172.16.254.226 (metric 20645) from 10.77.255.57 (10.77.255.57)
    Origin IGP, metric 0, localpref 100, valid, confed-external
    Extended Community: RT:1100:1001
    mpls labels in/out nolabel/362
(64955 65003) 65089
  172.16.254.226 (metric 20645) from 10.57.255.11 (10.57.255.11)
    Origin IGP, metric 0, localpref 100, valid, confed-external, best
    Extended Community: RT:1100:1001
    mpls labels in/out nolabel/362
!--- BGP selects this as the Best Path on comparing
!--- with all the other routes and selected based on lower router ID. (64955 65003) 65089
172.16.254.226 (metric 20645) from 172.16.224.253 (172.16.224.253) Origin IGP, metric 0,
localpref 100, valid, confed-internal Extended Community: RT:1100:1001 mpls labels in/out
nolabel/362 (65003) 65089 172.16.254.226 (metric 20645) from 172.16.254.234 (172.16.254.234)
Origin IGP, metric 0, localpref 100, valid, confed-external Extended Community: RT:1100:1001
mpls labels in/out nolabel/362 65089, (Received from a RR-client) 172.16.228.226 (metric 20645)
from 172.16.228.226 (172.16.228.226) Origin IGP, metric 0, localpref 100, valid, confed-internal
Extended Community: RT:1100:1001 mpls labels in/out nolabel/278
BGPはこのドキュメントで説明するさまざまな属性を考慮して、この9つのパスの中からベスト
パスを選択します。次に示す出力では、BGPが使用可能なパスを比較し、ルータIDが低いこと
からパス#6をベストパスとして選択しています。

Comparing path 1 with path 2:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
```

The paths have different neighbor AS's so ignoring MED
Both paths are internal
(no distinction is made between confed-internal and confed-external)
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 2 is better than path 1 because it has a lower Router-ID.

Comparing path 2 with path 3:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
Both paths have the same neighbor AS, 65089, so comparing MED.
Both paths have a MED of 0
Both paths are confed-external
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 2 is better than path 3 because it has a lower Router-ID.

Comparing path 2 with path 4:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
Both paths have the same neighbor AS, 65089, so comparing MED.
Both paths have a MED of 0
Both paths are confed-external
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 2 is better than path 4 because it has a lower Router-ID.

Comparing path 2 with path 5:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
Both paths have the same neighbor AS, 65089, so comparing MED.
Both paths have a MED of 0
Both paths are confed-external
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 5 is better than path 2 because it has a lower Router-ID.

Comparing path 5 with path 6:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
Both paths have the same neighbor AS, 65089, so comparing MED.
Both paths have a MED of 0
Both paths are confed-external
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 6 is better than path 5 because it has a lower Router-ID.

Comparing path 6 with path 7:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1

Both paths are of origin IGP
Both paths have the same neighbor AS, 65089, so comparing MED.
Both paths have a MED of 0
Both paths are internal
(no distinction is made between confed-internal and confed-external)
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 6 is better than path 7 because it has a lower Router-ID.

Comparing path 6 with path 8:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
Both paths have the same neighbor AS, 65089, so comparing MED.
Both paths have a MED of 0
Both paths are confed-external
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 6 is better than path 8 because it has a lower Router-ID.

Comparing path 6 with path 9:
Both paths have reachable next hops
Both paths have a WEIGHT of 0
Both paths have a LOCAL_PREF of 100
Both paths are learned
Both paths have AS_PATH length 1
Both paths are of origin IGP
The paths have different neighbor AS's so ignoring MED
Both paths are internal
(no distinction is made between confed-internal and confed-external)
Both paths have an IGP metric to the NEXT_HOP of 20645
Path 6 is better than path 9 because it has a lower Router-ID.
The best path is #6

パスを選択するプロセスのカスタマイズ

[BGP コスト コミュニティ](#)と呼ばれる拡張コミュニティ属性により、ベストパスの選択プロセスをカスタマイズすることができます。コストコミュニティが比較される追加のステップが、「[最適パスアルゴリズムの仕組み](#)」の項で説明されているアルゴリズムに追加されています。この手順は、アルゴリズムで要求される手順（挿入ポイント）の後に挿入されます。最小のコスト値を持つパスが優先されます。

注: 次の項目にも注意してください。

-このステップは [bgp bestpath cost-community ignore コマンド](#)を発行する場合スキップされます。

コミュニティ ID 番号 (255) への 0 をおよびコスト第 (0 に 4,294,967,295) 要するために -コスト コミュニティセツ節はで設定されます。コスト数値によってパスの優先度が判断されます。最小のコスト数値を持つパスが優先されます。具体的にコスト数値を使用して設定されているわけではないパスには、デフォルトのコスト数値である 2,147,483,647 が割り当てられます。この値は、0 ~ 4,294,967,295 の中央値です。次に、これらのパスは最適パス選択プロセスに従って評価されます。2つのパスが同じコスト数値を使用して設定されている場合、パス選択プロセスでは最も低いコミュニティ ID のパスが優先されます。パスの準最適パス コスト コミュニティが等しくない場合は、準最適パス コスト コミュニティの低いパスが最適パスとして選択されます。

- ABSOLUTE_VALUE は第一歩 パスのプリファレンスの次数の判別の考慮されます。たとえば、EIGRP が BGP VPNv4 に再配布されると、ABSOLUTE_VALUE タイプはコスト コミュニティに使用されます。IGB_Cost は、ネクスト ホップまでの内部 (IGP) 距離が比較された後に考慮されます。つまり、IGP_COST 挿入ポイントのコスト コミュニティが、「[最適パス アルゴリズムの仕組み](#)」のアルゴリズムの手順 8 の後で考慮されます。

BGP マルチパス

BGP マルチパスでは、同じ宛先への複数の BGP パスの IP ルーティング テーブルに対するインストールが許可されます。これらのパスは、ロードシェアリング用に最適パスとともにテーブル内へインストールされます。BGP マルチパスによる最適パスの選択への影響はありません。たとえば、ルータではアルゴリズムに従って、最適パスとしてパスの 1 つを引き続き指定しており、この最適パスを隣接ルータにアドバタイズします。

BGP マルチパスの機能には、次のものがあります。

- 多重通路 eBGP -[最大パス \$n\$](#)
- 多重通路 iBGP -[最大パス \$ibgp\ n\$](#)
- 多重通路 eiBGP -[最大パス \$eibgp\ n\$](#)

マルチパスの候補になるためには、同じ宛先へのパスに、最適パスの特性に等しいこれらの特性が備わっている必要があります。

- 重量
- Local preference
- AS-PATH length
- Origin
- MED
- 次のいずれか。ネイバー AS または sub-AS (eiBGP マルチパス機能が追加される前) AS-PATH (eiBGP マルチパス機能が追加された後)

一部の BGP マルチパス機能により、マルチパス候補に追加の要件が加わりました。

eBGP マルチパスには次の追加要件があります。

- パスは、外部またはコンフェデレーション外部の近接ルータ (eBGP) から学習されます。
- BGP ネクストホップへの IGP メトリックは、最適パスの IGP メトリックと等しくなる必要があります。

iBGP マルチパスには次の追加要件があります。

- パスは、内部の近接ルータ (iBGP) から学習されます。
- ルータが不等コスト iBGP マルチパスで設定されない限り、BGP ネクストホップへの IGP メトリックは、最適パスの IGP メトリックと等しくなる必要があります。

BGP は、マルチパス候補から最近受信した最大 n 個のパスを IP ルーティング テーブルに挿入します。現在、 n の最大値は 6 です。マルチパスがディセーブルになっている場合のデフォルト値は 1 です。

不等コスト ロード バランシングでは、BGP リンク帯域幅も使用できます。

注: 同等の next-hop-self は、内部ピアに転送される前に、eBGP マルチパス間で選択される

最適パスで実行されます。

関連情報

- [BGP に関するトラブルシューティング](#)
- [BGP ルータはベストパス選択でマルチエグジット識別子をどのように使うか](#)
- [BGP の設定](#)
- [BGP に関するサポート ページ](#)
- [テクニカルサポートとドキュメント - Cisco Systems](#)