Cisco Application Centric Infrastructure

# ACI の設定のベストプラクティス

このガイドは、ファブリックで有効にする必要がある ACI 設定を理解するために役立ちます。

## グローバル設定のベストプラクティス

**1. MCP を有効にする (VLAN ごとに):** MisCabling Protocol (MCP) は、外部ソース (誤動作しているサーバ、STP を実行している外部ネットワーク機器など) からのループを検出し、ACI が自身のパケットを受信するインターフェイスを err-disable にします。

1. これを有効にするには、[ ファブリック (Fabric) ] > [ アクセスポリシー (Access Policies) ] > [ グローバルポリシー (Global Policies) ] > [MCP インスタンスポリシーのデフォルト (MCP Instance Policy default) ] に移動します。
2. [VLAN ごとに MCP PDU を有効にする(Enable MCP PDU per VLAN)] オプション(2.0(2) 以降で利用可能) をオンにします。これにより、MCP は EPG ごとにパケットを送信できるようになります。それ以外の場合、これらのパケットはタグ設定のない EPG でのみ送信されます (これは、ループ検出の観点からは無用)。

**2. リモート EP 学習を無効にする** : これにより、ボーダーリーフスイッチでのリモート IP 学習が無効になります。

1. 3.0 より前の場合は、[ ファブリック (Fabric) ] > [ アクセスポリシー (Access Policies) ] > [ グローバルポリシー (Global Policies) ] > [ ファブリック全体の設定ポリシー (Fabric Wide Setting Policy) ] に移動すると、これを有効にすることができます。
2. 3.0 以降の場合は、[ システム (System) ] > [ システム設定 (System Settings) ] > [ ファブリック全体の設定 (Fabric Wide Setting) ] に移動すると、これを有効にすることができます。
3. これは、2.2(2e) 以降のすべてのコードで利用可能です。
4. *CSCvi11291 に注意してください (3.2(1l) 以降では修正済み)。このバグにより、[ リモート EP 学習の無効化 (Disable Remote EP Learn)] がオンになっている場合でも、送信先 / 宛先が tcp 179 のパケットをスイッチが受信した場合のボーダーリーフスイッチでのリモート EP 学習が可能になります。*

**3. サブネットチェックを適用する** : これは、-EX および -FX ベースのリーフでのみ機能します。

1. 3.0 より前の場合は、[ ファブリック (Fabric) ] > [ アクセスポリシー (Access Policies) ] > [ グローバルポリシー (Global Policies) ] > [ ファブリック全体の設定ポリシー (Fabric Wide Setting Policy) ] に移動すると、これを有効にすることができます。
2. 3.0 以降の場合は、[ システム (System) ] > [ システム設定 (System Settings) ] > [ ファブリック全体の設定 (Fabric Wide Setting) ] に移動すると、これを有効にすることができます。
3. これは、2.2(2q) 以降のすべての 2.2(x) コードで利用可能です。
    1. 2.3(x) では利用できません。
    2. 3.0 (3.0(2k) 以降) から利用できます。
4. [ サブネットチェックの適用 (Enforce Subnet Check) ] は [IP 学習をサブネットに限定 (Limit IP Learning to subnet) ] に似ています。ただし、BD 設定オプションの [IP 学習をサブネットに限定 (Limit IP Learning to subnet) ] は、BD で設定されたサブネットに含まれない IP エンドポイントの学習を防止することに注意してください。[IP 学習をサブネットに限定 (Limit IP Learning to subnet) ] ではパケットはドロップされません。BD で学習が設定されなくなるだけです。BD が設定されていないリーフ (つまり、ボーダーリーフ) では引き続きパケットの学習が可能です。これによって問題が発生する可能性があるため、[ サブネットチェックの適用 (Enforce Subnet Check) ] 設定オプションが必要になります。オンにすると、VRF レベルでも IP コンポーネントが学習されなくなります。

5. *CSCvh17285 に注意してください (3.2(1l) 以降では修正済み)。[ サブネットチェックの適用 (Enforce Subnet Check) ] がオンになっている場合、L2 のみとして設定されたブリッジドメインで、L2 Unknown ユニキャストがプロキシに設定されていると、MAC アドレスが ARP/GARP パケットから学習されません。回避策は、L2 BD の L2 Unknown ユニキャスト設定をフラッドにすることです。*

### 4. EP ループ検出を設定する

1. EP ループ検出設定には適切な目的 (ループを検出して終了させる) がありますが、誤検出 (VM の Vmotion など) によってトリガーされることも少なくありません。そのため、この設定は有効のままにしておくことが推奨されるものの、両方のアクション (BD 学習の無効化とポートの無効化) が無効 (オフ) になっていることを確認してください。両方のアクションを無効にしても、引き続き EP ループによってエラーが生成され、Syslog/SNMP トラップサーバに送信されます (設定されている場合)。
2. 3.0 より前の場合は、[ ファブリック (Fabric) ] > [ アクセスポリシー (Access Policies) ] > [ グローバルポリシー (Global Policies) ] > [EP ループ検出ポリシー (EP Loop Detection Policy) ] に移動すると、これを有効 (または無効) にすることができます。
3. 3.0 以降の場合は、[ システム (System) ] > [ システム設定 (System Settings) ] > [ エンドポイント制御 (Endpoint Controls) ] > [EP ループ検出 (EP Loop Detection) ] に移動すると、これを有効にすることができます。

### 5. IP エージングが有効になっていることを確認する

1. IP エージングが有効になっていない (デフォルト) 場合、単一の MAC で複数の IP が学習されると、その MAC がアクティブであるかぎり、すべての IP がファブリック上で学習されたままになります。表面的には、これは、DHCP 対応ホストが新しい IP アドレスを取得したのに、両方の IP がその MAC に関連付けられているものとして EPG 動作タブ内に表示されるようなシナリオにおいて、望ましくありません。IP エージング機能により、そのようなシナリオに対処するために、各 IP が個別にエージングされます。エンドポイント保持タイマーが 75% で、ダイレクトされた ARP がエンドポイントの IP コンポーネントに送信され、応答がない場合、ACI は IP エンドポイントのエージアウトを許可します。
2. 3.0 より前の場合は、[ ファブリック (Fabric) ] > [ アクセスポリシー (Access Policies) ] > [ グローバルポリシー (Global Policies) ] > [IP エージングポリシー (IP Aging Policy) ] に移動すると、これを有効にすることができます。
3. 3.0 以降の場合は、[ システム (System) ] > [ システム設定 (System Settings) ] > [ エンドポイント制御 (Endpoint Control) ] > [IP エージング (IP Aging) ] に移動すると、これを有効にすることができます (このタブの右側を参照)。
4. これは、2.1(1h) 以降のすべてのコードで利用可能です。

### 6. 不正エンドポイント検出が有効になっていることを確認する

1. 3.2 以降の場合は、不正エンドポイントの検出により、エンドポイントのフラッピングによる影響が軽減されます。
2. 不正エンドポイント検出が有効になっている場合、動作が不審なエンドポイント (MAC/IP) が隔離され、簡単に識別できるようにエラーが生成されます。
3. 3.2 以降の場合は、[ システム (System) ] > [ システム設定 (System Settings) ] > [ エンドポイント制御 (Endpoint Controls) ] > [ 不正 EP 制御 (Rogue EP Control) ] に移動すると、これを有効にすることができます。
4. 推奨値は、次のとおりです。
    1. 不正 EP 検出の間隔 = 30
    2. 不正 EP 検出の乗算係数 = 6

### 7. 厳密な COOP グループポリシーを有効にする

1. APIC は、MD5 パスワードに使用される属性を含む管理対象オブジェクト (fabric: SecurityToken) を提供します。この管理対象オブジェクト内の属性 (「トークン」と呼ばれる) は、1 時間ごとに変更される文字列です。COOP は、DME から通知を受け取り、ZMQ 認証のパスワードを更新します。この属性トークンの値は表示されません。互換性タイプと厳密タイプの 2 つの選択肢があります。互換性タイプでは MD5 認証済み ZMQ 接続と MD5 非認証 ZMQ 接続の両方が許可されますが、厳密タイプでは MD5 認証済み ZMQ 接続だけが許可されます。
2. [ システム (System) ] > [ システム設定 (System Settings) ] > [COOP グループ (COOP Group) ] に移動すると、これを有効にすることができます。

**8. ファブリックに接続しているインターフェイスの BFD を有効にする**

1. ファブリック間インターフェイス（リーフからスパイン）の BFD により、障害発生時のコンバージェンスが高速化されます。
2. [ ファブリック (Fabric) ] > [ ファブリックポリシー (Fabric Policies) ] > [ ポリシー (Policies) ] > [L3 インターフェイス (L3 Interface) ] > [ デフォルト (default) ] > [BFD ISIS ポリシー設定 (BFD ISIS Policy Configuration) ] に移動すると、これを有効にすることができます。

**9. ACI ファブリックを通過する CoS を保持する**

1. [ ファブリック (Fabric) ] > [ アクセスポリシー (Access Policies) ] > [ ポリシー (Policies) ] > [ グローバル (Global) ] > [QOS クラス (QOS Class) ] > [COS の保持 (Preserve COS) ] に移動すると、これを有効にすることができます。
2. APIC により、ファブリック内で 802.1P サービスクラス (CoS) 設定を保持することができます。ファブリックグローバル QoS ポリシーの dot1p-preserve オプションを有効にすることで、ACI ファブリックに出入りするパケットの CoS 値の保持が保証されます。802.1P CoS の保持は、シングルポッドトポロジとマルチポッドトポロジでサポートされます。

大まかに言うと、オプション 2 とオプション 3 は、発生する可能性のあるファブリック上の IP エンドポイントの誤学習を防止します。エンドポイントの誤学習により、リモート IP エンドポイントがボーダーリーフでスタックする可能性があるため、パケットのブラックホールのようなものの原因になります。このようなイベントを解消するプロセスは単純ではありません。各エンドポイント設定オプションの使用例の詳細については、ACI ファブリックのエンドポイントの学習に関する次のホワイトペーパーを参照してください。

### 『ACI Fabric Endpoint Learning』ホワイトペーパー

https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.pdf

## ブリッジドメインのベストプラクティス：

ブリッジドメインには、さまざまな使用例があります。ベストプラクティスは、状況によって異なります。ただし、以下のように、いくつかの包括的な推奨事項と関連する注意事項があります。

**1. ACI がサブネットの L3 ゲートウェイでない場合は、ユニキャストルーティングを有効にしないでください。**ACI が L3 ゲートウェイでない場合にユニキャストルーティングを有効にするのは、有効にしないと ACI がエンドポイントの IP アドレスを学習しないからです。そのため、接続デバイスの MAC アドレスだけではなく IP エンドポイントも学習する必要がある場合は、当然のことながら、ユニキャストルーティングを有効にすることになります。ただし、これには、非対称ルーティングの原因になり、その結果として、パケットがドロップされたり誤ってルーティングされる可能性があるという問題があります。

**2. ブリッジドメインごとに単一のサブネットを設定してください。**ACI は、各 BD のプライマリサブネットでのみ DHCP 要求を転送します。同じ BD に 2 つ目のサブネットが設定されている場合、DHCP は 2 つ目以降の BD では機能しません。

**3.** ネットワークセントリック モード（つまり、VLAN=EPG=BD）の **BD に複数の EPG を設定しないでください。**ACI で VLAN を EPG と BD にマッピングする場合、外部 STP および HSRP マルチキャストは同じ BD にフラッディングされます。たとえば、Vlan11 (EPG11) と Vlan12 (EPG12) が同じ BD に接続されている場合、両方の VLAN の HSRP hello が BD 内で混在し、外部（非 ACI）環境で問題が発生します。

4. **[IP 学習をサブネットに限定 (Limit IP Learning to Subnet) ] を有効にしてください。**これは、ほとんどの場合に有効にする必要があります。これにより、エンドポイントの IP 学習が、各ブリッジドメインで設定されているサブネットに基づいて制限されます。注 : -EX または -FX ベースのリーフがあり、[ サブネットチェックの適用 (Enforce Subnet Check) ] をグローバルに設定している場合、これは自動的にオンになります。

5. **ARP フラッディング + GARP ベースの検出を検討してください。**この設定は、一部の環境では実現されていますが、選択されていない環境もあります。GARP ベースの検出の長所は、特定の状況で IP 学習に関する問題が防止されることです。短所は、GARP ベースの検出を設定する前に BD で ARP フラッディングを有効にする必要があることです。

『Cisco ACI Fabric Endpoint Learning』ホワイトペーパーでは、次のように説明されています。
「Cisco ACI は、リーフスイッチポート、リーフスイッチ、ブリッジドメイン、および EPG の間での MAC および IP アドレスの移動を検出できますが、新しい MAC アドレスが古い MAC アドレスと同じインターフェイスおよび同じ EPG からのものである場合、その新しい MAC アドレスへの IP アドレスの移動を検出しません。GARP ベースの検出のオプションが有効になっている場合、同じインターフェイスおよび同じ EPG での移動が発生すると、Cisco ACI は GARP パケットに基づいてエンドポイントの移動をトリガーします。GARP パケットが同じインターフェイスおよび同じ EPG から着信すると、ユニキャストルーティング、ARP フラッディング、および「GARP ベースの検出」のすべてがブリッジドメインで有効になっている場合にのみエンドポイント学習がトリガーされます。このシナリオは、シスコの顧客基盤全体で広く見られるものではありませんが、場合によっては、IP から MAC へのバインディングを変更し、GARP ベースの検出を有効にする必要があります」

# ファブリック プロビジョニングのベストプラクティス

ACI ファブリックのセットアップを実行するには、セットアップ値の適切なプランニングが不可欠です。*ACI ファブリックの値を検討する際は、「初期プロビジョニング セットアップ プロセスの完了後は、ファブリックを再構築しないと、インフラストラクチャ IP アドレス (TEP IP プール) 範囲またはインフラストラクチャ VLAN のいずれかを変更することはできない」ことを覚えておくことが重要です。*

ファブリックの初期セットアップを実行するときは、「TEP アドレス範囲」を入力する必要があります。この範囲の IP アドレスは、主に、ファブリック内のリーフノードとスパインノードに TEP アドレスを提供するために使用されます。これのデフォルト値は 10.0.0.0/16 ですが、次のようないくつかの理由から、TEP プールに一意のアドレスブロックを提供することがベストプラクティスであると見なされます。

1. 将来、TEP プールを AVE (ACI Virtual Edge) スイッチに拡張する場合は、ネットワーク内の既存のルーティングと重複しない一意のアドレスを作成する必要があります。
2. APIC から外部デバイス(つまりVMM 統合のための vCenter) と通信する場合vCenter デバイスから APIC に戻るトラフィックの IP アドレス / ルーティングの競合を回避するために、インフラストラクチャ TEP プールで一意のアドレッシングを作成する必要があります。
3. 注 : 初期プロビジョニング後にインフラストラクチャ IP アドレス範囲または VLAN を変更するには、ファブリックを再構築する必要があります。

インフラストラクチャ サブネットは、ネットワーク内の他のルーティングされるサブネットと重複しないようにする必要があります。このサブネットが別のサブネットと重複している場合は、このサブネットを別の /16 サブネットに変更してください。

- APIC 2.2 コード以降の場合、3-APIC クラスタでサポートされる最小サブネットは /23 です。
- APIC 2.2 コードまでの APIC 2.0(1) コードを使用している場合、最小サブネットは /22 です。
- インフラストラクチャ TEP IP は、未使用で一意である必要があります。ただし、予備の RFC1918 アドレスがない場合は、RFC6598 範囲 (100.64/10 : CGN 用) の使用を検討してください。それにより、インターネット上で競合しないことが保証されます。
- すべてのファブリック /POD インフラストラクチャ TEP プールは、一意の IP サブネット範囲から作成される必要があります。

詳細については、『Cisco APIC Getting Started Guide, Release 3.x』を参照してください。

**インフラストラクチャ VLAN ID** : インフラストラクチャ VLAN を 3967 に設定してください。

ファブリックのセットアップ時に、ACI は、インフラストラクチャ VLAN として使用する VLAN を要求します。この VLAN は、ファブリックを構成するデバイス（つまり、リーフ、スパイン、および APIC）間のトラフィックを制御するために使用されます。

この VLAN はファブリックの外部に拡張できるため（Openstack 統合、AVS/AVE）、これを環境内で一意の VLAN として持つことがベストプラクティスとなります。さらに、シスコの多くのデバイスでは、変更が難しい VLAN 範囲が予約されています（変更を有効にするにはスイッチの再起動が必要）。VLAN 3967 は、シスコのスイッチング プラットフォームで予約されておらず、ACI に最適な VLAN です。

**ノード ID 設定** : スパインには 101 ～ 199 の番号を割り当て、リーフには 200 以上の番号を割り当てる必要があります。

詳細については、『Cisco Application Centric Infrastructure Best Practices Guide』の『Fabric Provisioning』を参照してください。

# ACI ファブリックの命名のベストプラクティス

ACI ファブリックの命名のベストプラクティスについては、優れた入門書があります。こちらの投稿で、ファブリックのテナントセクションとファブリックアクセスセクションの両方におけるオブジェクト命名に関するヒントを確認してください。