ICMPリダイレクトメッセージについて

内容

概要

このドキュメントでは、インターネット制御メッセージプロトコル(ICMP)パケットリダイレクト機能について説明します。

前提条件

要件

次の項目に関する知識があることが推奨されます。

- Nexus 7000プラットフォーム アーキテクチャ
- Cisco NX-OSソフトウェアの設定
- Request for Comments（RFC）792 で規定されている Internet Control Message Protocol

使用するコンポーネント

このドキュメントの情報は、次のソフトウェアとハードウェアのバージョンに基づいています。

- Nexus 7000
- Cisco NX-OS ソフトウェア

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、初期（デフォルト）設定の状態から起動しています。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してく

ださい。

背景説明

このドキュメントは、Internet Control Message Protocol（ICMP）によって提供されるパケットリダイレクト機能に関するものです。 ネットワーク内に ICMP リダイレクトメッセージが存在することは、通常、何を意味しているのか、また、ICMP リダイレクトメッセージの生成を引き起こしたネットワークの状態に関連する悪影響を最小限に抑えるために何ができるのかを説明します。

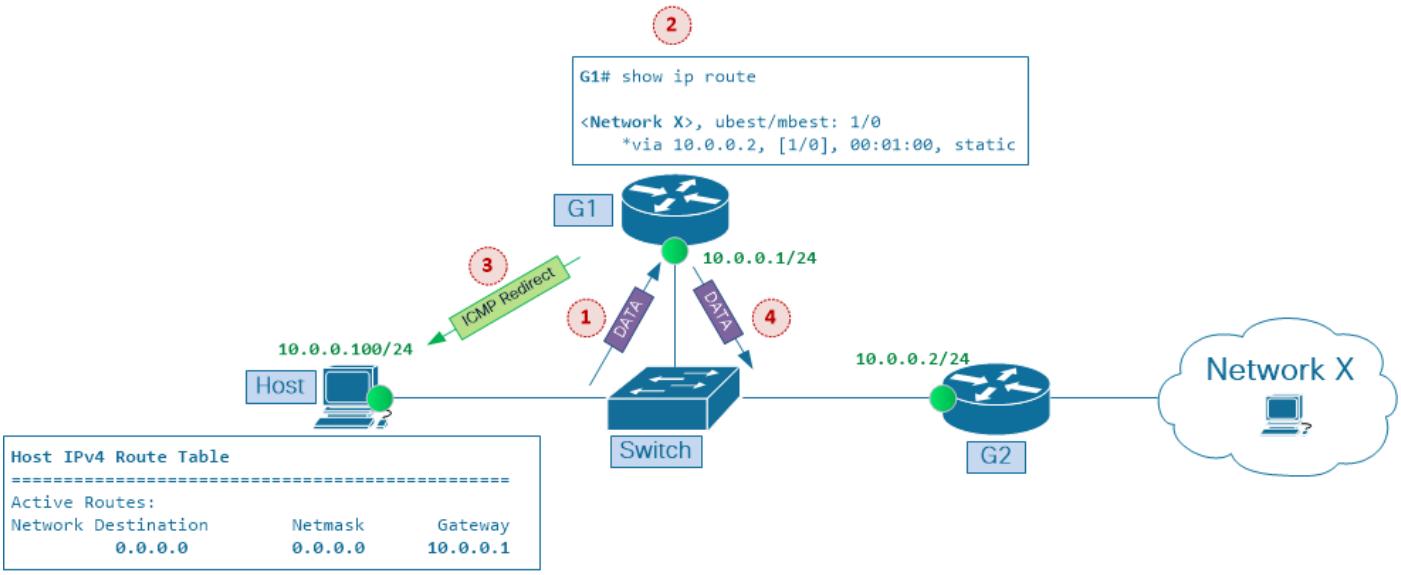
ICMP リダイレクトメッセージ

ICMPリダイレクト機能は、次の例を使用して[RFC 792 Internet Control Message Protocol]で説明されています。

> この場合、ゲートウェイはホストにリダイレクトメッセージを送信します。
>
> ゲートウェイG1は、ゲートウェイが接続されているネットワーク上のホストからインターネットデータグラムを受信する。ゲートウェイG1は自身のルーティングテーブルをチェックし、データグラムのインターネット宛先ネットワークXへのルート上の次のゲートウェイG2のアドレスを取得します
>
> G2と、データグラムのインターネット送信元アドレスによって識別されるホストが同じネットワーク上にある場合、リダイレクトメッセージがホストに送信されます。リダイレクトメッセージはホストに対して、ネットワーク X のトラフィックをゲートウェイ G2 に直接送信することを通知します（宛先までのより短いパスであるため）。
>
> ゲートウェイは、元のデータグラムデータをインターネット宛先に転送します。

このシナリオを図1に示します。ホストと2台のルータ（G1およびG2）が共有イーサネットセグメントに接続され、同じネットワーク10.0.0.0/24にIPアドレスを持っています

図1：マルチポイントイーサネットネットワークでのICMPリダイレクト



マルチポイントイーサネットネットワークでのICMPリダイレクト

ホストにはIPアドレス10.0.0.100があります。ホストルーティングテーブルには、ルータG1のIPアドレス10.0.0.1をデフォルトゲートウェイとして指すデフォルトルートエントリがあります。ルータ G1 は、宛先ネットワーク X にトラフィックを転送するときに、ルータ G2 の IP アドレス（10.0.0.2）をネクストホップとして使用します。

ホストが宛先ネットワークXにパケットを送信すると、次のようになります。

1. IPアドレス10.0.0.1のゲートウェイG1は、接続先のネットワーク上のホスト10.0.0.100からデータパケットを受信します。

2.ゲートウェイG1は、自身のルーティングテーブルを確認し、データパケットの宛先ネットワークXへのルート上の次のゲートウェイG2のIPアドレス10.0.0.2を取得します。

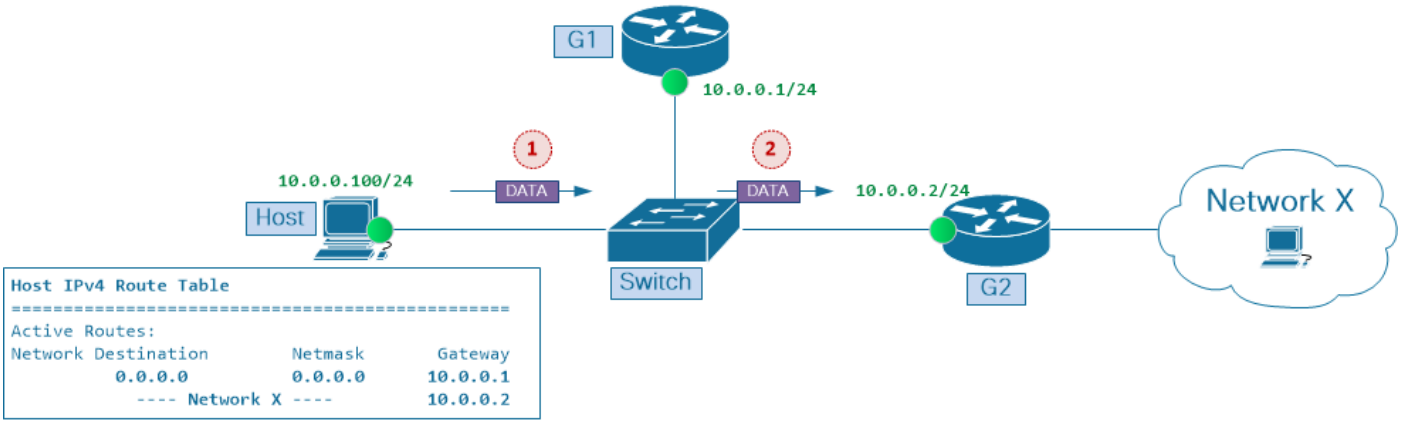
3. G2とIPパケットの送信元アドレスで識別されるホストが同じネットワーク上にある場合、ICMPリダイレクトメッセージがそのホストに送信されます。ICMPリダイレクトメッセージは、宛先へのパスが短いため、ネットワークXへのトラフィックをゲートウェイG2に直接送信するようにホストに通知します。

4. ゲートウェイ G1 が、元のデータパケットを宛先に転送します。

ホストの設定に応じて、G1が送信するICMPリダイレクトメッセージを無視することを選択できます。ただし、ホストがICMPリダイレクトメッセージを使用してルーティングキャッシュを調整し、後続のデータパケットをG2に直接送信し始めた場合は、このシナリオで次の利点が得られます

- ネットワーク経由のデータ転送パスの最適化トラフィックは宛先に高速に到達する
- 帯域幅やルータの CPU 負荷といったネットワークリソースの使用率が削減されます。

図2ホストルーティングキャッシュにインストールされたネクストホップG2



ホストルーティングキャッシュにインストールされたネクストホップG2

図2に示すように、ホストがG2をネクストホップとしてネットワークXのルートキャッシュエントリを作成した後、ネットワークには次のような利点があります。

- スイッチとルータG1間のリンクの帯域幅使用率は、両方向で低下します。
- ホストからネットワークXへのトラフィックフローがこのノードを通過しなくなるため、ルータG1のCPU使用率が低下します。
- ホストとネットワーク X の間のエンドツーエンドのネットワーク遅延が改善されます。

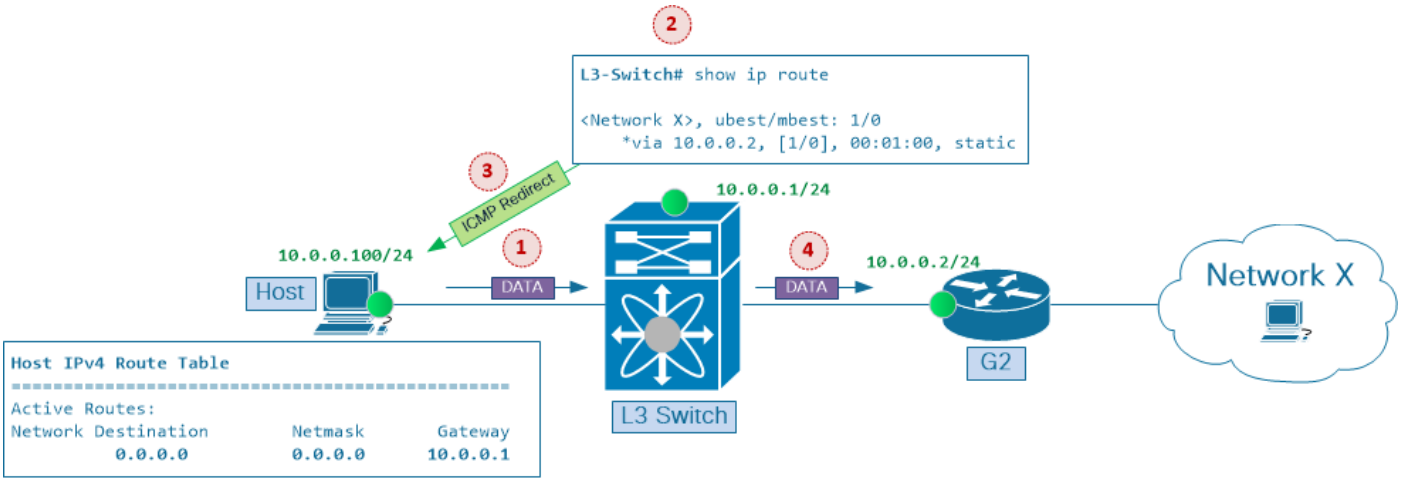
ICMP リダイレクトメカニズムの重要性を理解するには、初期のインターネットルータ実装がデ

ータトラフィックを処理するために主に CPU リソースに依存していたことを思い出す必要があります。したがって、任意の1台のルータで処理する必要があるトラフィック量を減らし、特定のトラフィックフローが宛先に向かう途中で通過しなければならないルータホップ数を最小限に抑えることが望ましくなりました。同時に、レイヤ 2 転送（スイッチングとも呼ばれる）は主にカスタマイズされた特定用途向け集積回路（ASIC）で実装され、転送パフォーマンスの観点からは、やはり汎用プロセッサで実装されるレイヤ 3 転送（ルーティングとも呼ばれる）よりも比較的「貧弱」でした。

より新しい ASIC 世代は、レイヤ 2 とレイヤ 3 の両方のパケット転送を実行できます。ハードウェアで実行されるレイヤ3テーブルルックアップは、ルータによるパケット処理に関連するパフォーマンスコストを削減するのに役立ちます。さらに、レイヤ2スイッチにレイヤ3転送機能を統合（現在はレイヤ3スイッチと呼ばれている）してパケット転送動作を効率化すると、**ワンアームルータ**(router on a stickとも呼ばれる)の設計オプションが不要になり、このようなネットワーク設定に関連する制限を回避できます。

図3は、図1のシナリオに基づいて作成されています。現在、レイヤ2とレイヤ3の機能は、元々は2つの個別のノードであるスイッチとルータG1によって提供されていましたが、Nexus 7000シリーズプラットフォームなどの1つのレイヤ3スイッチに統合されています。

図3 「ワンアームドルータ」の設定を置き換えるレイヤ3スイッチ



「One-armed-router」設定をレイヤ3スイッチに置き換え

ホストが宛先ネットワークXにパケットを送信すると、次のようになります。

1. IPアドレスが10.0.0.1のゲートウェイL3スイッチが、接続されているネットワーク上のホスト10.0.0.100からデータパケットを受信する。

2.ゲートウェイL3スイッチは、自身のルーティングテーブルをチェックし、データパケット宛先ネットワークXへのルート上の次のゲートウェイG2のアドレス10.0.0.2を取得します。

3. G2とIPパケットの送信元アドレスで識別されるホストが同じネットワーク上にある場合、ICMPリダイレクトメッセージがそのホストに送信されます。ICMPリダイレクトメッセージは、宛先へのパスが短いため、ネットワークXへのトラフィックをゲートウェイG2に直接送信するようにホストに通知します。

4. ゲートウェイが、元のデータパケットを宛先に転送します。

レイヤ3スイッチがASICレベルでレイヤ2とレイヤ3の両方のパケット転送を実行できるようになったことで、ICMPリダイレクト機能の利点、(a)ネットワーク遅延の改善、および(b)ネットワー

クリソース使用率の削減の両方が達成され、マルチポイントイーサネットセグメントでのパス最適化技術に対する注意を払う必要がなくなったと結論付けることができます。

ただし、レイヤ 3 インターフェイスで ICMP リダイレクト機能が有効になっていると、別の理由から、マルチポイント イーサネット セグメントを介した準最適転送によって引き続きパフォーマンスボトルネックが発生する可能性があります（このドキュメントの「Nexusプラットフォームに関する考慮事項」を参照）。

注：ICMPリダイレクトは、Cisco IOSおよびCisco NX-OSソフトウェアのレイヤ3インターフェイスでデフォルトで有効になっています。

注：ICMP リダイレクトメッセージが生成される条件の概要：レイヤ 3 スイッチは、データパケットが受信されたレイヤ 3 インターフェイスからそのパケットが転送される場合に、そのパケットの送信元に対して ICMP リダイレクトメッセージを生成して返します。

イーサネットネットワークを介した準最適パス

Open Shortest Path First（OSFP）や Cisco Enhanced Interior Gateway Routing Protocol（EIGRP）などの内部ゲートウェイプロトコル（IGP）は、ルータ間でルーティング情報を同期させ、そのような情報を受け取るすべてのネットワークノードで一貫性のある予測可能なパケット転送動作を実現するように設計されています。たとえば、マルチポイントイーサネットネットワークでは、セグメント上のすべてのレイヤ3ノードが同じルーティング情報を使用し、宛先への同じ出力点で合意した場合、このようなネットワークを経由する最適ではない転送はめったにありません。

準最適転送パスの原因を理解するには、レイヤ 3 ノードがパケット転送の決定を相互に独立して行うことを思い出す必要があります。つまり、ルータ B によるパケット転送の決定は、ルータ A によるパケット転送の決定に依存しません。これは、IPネットワークを介したパケット転送のトラブルシューティングを行う際に覚えておくべき重要な原則の1つであり、マルチポイントイーサネットネットワークで最適ではない転送パスを調べる際に注意する必要があります。

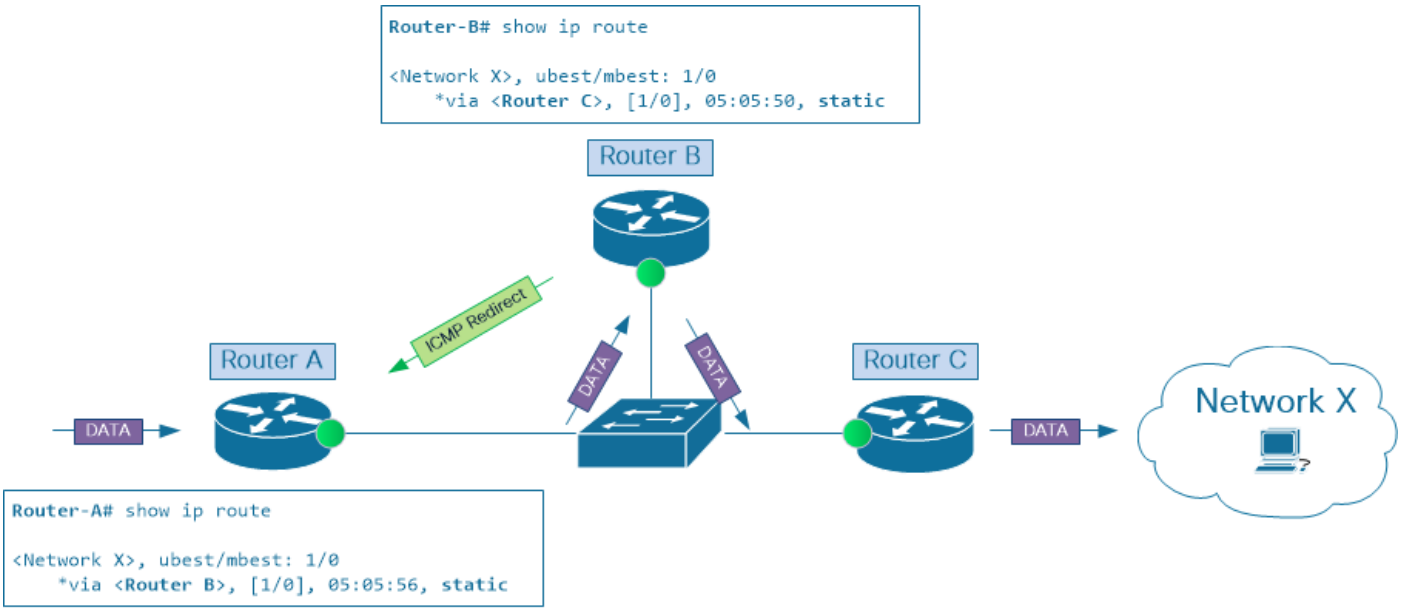
前述したように、すべてのルータが単一のダイナミックルーティングプロトコルを使用してエンドポイント間のトラフィックを配信するネットワークでは、マルチポイントイーサネットセグメントを介した最適でない転送は行われてはなりません。しかし、実際のネットワークでは、さまざまなパケットルーティングおよび転送メカニズムの組み合わせを見つけることが非常に一般的です。このようなメカニズムの例としては、さまざまなIGP、スタティックルーティング、ポリシーベースルーティングがあります。これらの機能は、通常、ネットワークを介して必要なトラフィック転送を行うために併用されます。

これらのメカニズムを組み合わせて使用すると、トラフィックフローを微調整し、特定のネットワーク設計の要件を満たすことができます。ただし、これらのツールを組み合わせて使用すると、マルチポイントイーサネットネットワークで発生する可能性がある副作用が見過ごされ、ネットワーク全体のパフォーマンスが低下する可能性があります。

スタティックルーティング

これを説明するために、図4のシナリオを考えてみましょう。ルータAには、ネクストホップとしてルータBを使用するネットワークXへのスタティックルートがあります。同時に、ルータ B は、ネットワーク X へのスタティックルートでルータ C をネクストホップとして使用します。

図4スタティックルーティングによる最適でないパス



スタティックルーティングによる最適でないパス

トラフィックは、ルータ A からこのネットワークに入り、ルータ C を通過して、最終的に宛先ネットワーク X に配信されますが、パケットは、宛先に到達するまでにこの IP ネットワークを 2 回通過する必要があります。これはネットワークリソースの効率的な使用ではありません。代わりに、ルータAからルータCに直接パケットを送信しても同じ結果が得られますが、ネットワークリソースの消費は少なくなります。

> **注**：このシナリオでは、ルータAとルータCがこのIPネットワークセグメントの入力および出力レイヤ3ノードとして使用されていますが、両方のノードに同じパケット転送動作を引き起こすルーティング設定がある場合は、両方のノードをネットワークアプライアンス（ロードバランサやファイアウォールなど）に置き換えることができます。

ポリシー ベース ルーティング

ポリシーベースルーティング（PBR）は、イーサネットネットワークを介した準最適パスを発生させる可能性があるもう一つのメカニズムです。ただし、スタティックルーティングやダイナミックルーティングとは異なり、PBRはルーティングテーブルレベルでは動作しません。その代わりに、トラフィックリダイレクト用のアクセスコントロールリスト(ACL)をスイッチハードウェアで直接処理します。その結果、選択されたトラフィックフローに対して、入力ラインカードでのパケット転送検索では、スタティックルーティングやダイナミックルーティングで取得されたルーティング情報がバイパスされます。

図4では、ルータAとルータBは宛先ネットワークXに関するルーティング情報をダイナミックルーティングプロトコルの1つと交換します。ルータBはこのネットワークへの最適なネクストホップです。

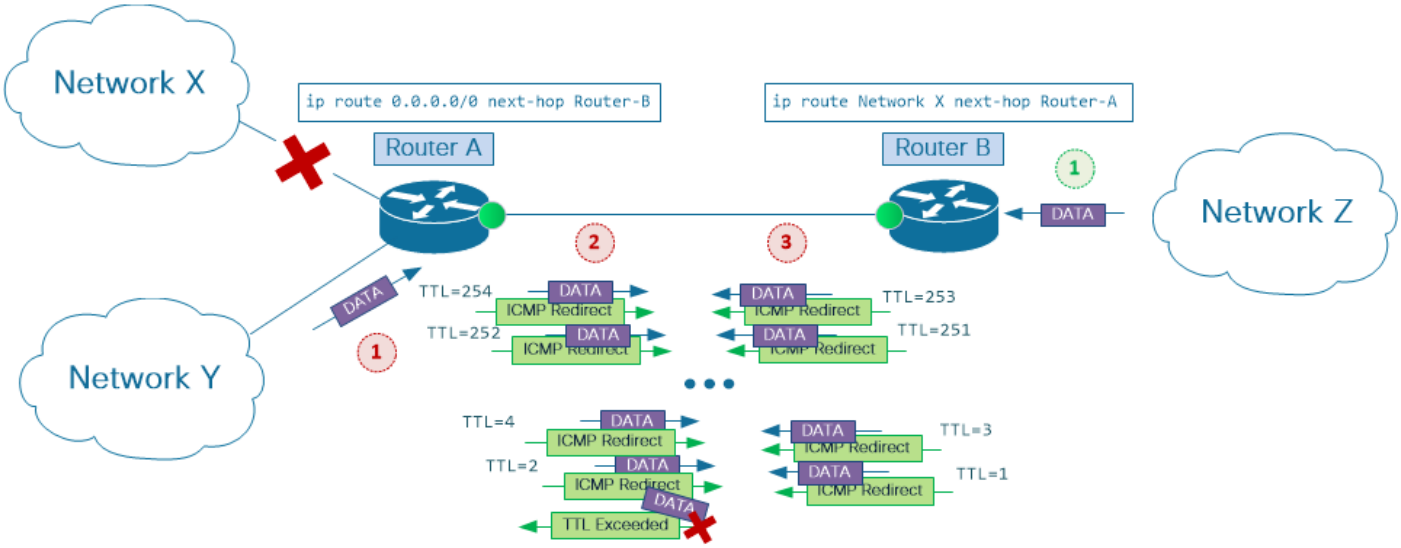
ただし、ルーティングプロトコルから受信したルーティング情報を上書きし、Router Cをネットワークxへのネクストホップとして設定するRouter B上のPBR設定では、ICMPリダイレクト機能をトリガーする条件が満たされると、パケットはRouter BのCPUに送信されてさらに処理されます。

ポイントツーポイントリンクでの ICMP リダイレクト

これまでのところ、このドキュメントでは、3 つ (またはそれ以上) のレイヤ 3 ノードが接続されたイーサネットネットワーク (そのため、マルチポイント イーサネット ネットワークと呼ばれる) について説明してきました。ただし、ICMP リダイレクトメッセージはポイントツーポイント イーサネット リンクでも生成される可能性があることに注意してください。

図5のシナリオを考えてみます。ルータAはスタティックデフォルトルートを使用してトラフィックをルータBに送信しますが、ルータBにはルータAをポイントするネットワークXへのスタティックルートがあります。

図5：ポイントツーポイントリンクでのICMPリダイレクト



スタティックルーティングによる最適でないパス

この設計オプションはシングルホーム接続とも呼ばれ、小規模なユーザ環境をサービスプロバイダーネットワークに接続する際に一般的な選択肢となります。ここで、ルータBはプロバイダーエッジ(PE)デバイスで、ルータAはユーザエッジ(CE)デバイスです。

一般的なCE設定には、Null0インターフェイスをポイントするユーザIPアドレスブロックへの集約スタティックルートが含まれていることに注意してください。この設定は、スタティックルーティングを使用したシングルホーム CE-PE 接続オプションの推奨ベストプラクティスです。ただし、この例では、そのような設定が存在しないものとします。

図に示すように、ルータAがネットワークXへの接続を失うと仮定します。ユーザのネットワークYまたはリモートのネットワークZからのパケットがネットワークXに到達しようとすると、ルータAとBは互いにトラフィックをバウンスでき、その値が1に達するまで各パケットのIP Time-To-Live (TTL ; 存続可能時間) フィールドを減少させます。この時点では、パケットのルーティングはこれ以上行えません。

ネットワークXへのトラフィックはPEルータとCEルータの間で往復しますが、CE-PEリンクの帯域幅使用率は大幅に (不必要に) 増加します。この問題は、ポイントツーポイントPE-CE接続の片側または両側でICMPリダイレクトが有効になっている場合に悪化します。この場合、ネットワークXに向けられたフローのすべてのパケットは、ICMPリダイレクトメッセージの生成に役立てるために、各ルータのCPUで複数回処理されます。

Nexus プラットフォームに関する考慮事項

ICMPリダイレクトがレイヤ3インターフェイスで有効になっていて、着信データパケットがこのインターフェイスを使用してレイヤ3スイッチの入力と出力の両方を行う場合、ICMPリダイレクトメッセージが生成されます。レイヤ3パケット転送はCisco Nexus 7000プラットフォームのハードウェアで行われますが、ICMPリダイレクトメッセージを作成するのはスイッチCPUの責任です。これは、入力ラインカードからスーパーバイザモジュールにデータパケットが送信される理由です。

ICMPリダイレクトメッセージの受信者がこのメッセージを無視し、ICMPリダイレクトが有効になっているNexusスイッチのレイヤ3インターフェイスにデータトラフィックを転送し続けると、各データパケットに対してICMPリダイレクト生成プロセスがトリガーされます。

ラインカードレベルでは、ハードウェア転送例外の形式でプロセスが開始されます。パケット転送操作がラインカードモジュールによって正常に完了できない場合、ASICで例外が発生します。この場合、データパケットを正常に処理するには、パケットがスーパーバイザモジュールに送信される必要があります。

> **注**：スーパーバイザモジュールのCPUは、ICMPリダイレクトメッセージを生成するだけでなく、存続可能時間(TTL)値が1に設定されたIPパケットや、ネクストホップに送信される前にフラグメント化される必要があるIPパケットなど、他の多くのパケット転送例外を処理します。

スーパーバイザモジュールのCPUがICMPリダイレクトメッセージを送信元に送信した後、出力ラインカードモジュールを介してデータパケットをネクストホップに転送することで、例外処理を完了します。

Nexus 7000スーパーバイザモジュールは、大量のトラフィックを処理できる強力なCPUプロセッサを使用しますが、このプラットフォームは、スーパーバイザCPUプロセッサをパケット転送処理に関与させることなく、ほとんどのデータトラフィックをラインカードレベルで処理するように設計されています。これにより、CPUはコアタスクに集中でき、パケット転送操作はラインカード上の専用ハードウェアエンジンに任されます。

安定したネットワークでは、パケット転送の例外が発生した場合、それなりに低いレートで発生することが予想されます。この前提に基づいて、そのパフォーマンスに大きな影響を与えることなく、スーパーバイザCPUでこれらの例外を処理できます。一方、非常に高いレートで発生するパケット転送例外を処理するCPUでは、システム全体の安定性と応答性に悪影響を及ぼす可能性があります。

Nexus 7000プラットフォームの設計には、スイッチのCPUを大量のトラフィックから保護するための多数のメカニズムが備わっています。これらのメカニズムは、システムのさまざまなポイントで実装されます。ラインカードレベルには、ハードウェアレートリミッタとコントロールプレーンがあります Policing (CoPP)機能。どちらもトラフィックレートのしきい値を設定します。これにより、各ラインカードモジュールからスーパーバイザに転送されるトラフィックの量が効果的に制御されます。

これらの保護メカニズムは、ネットワークの安定性やスイッチの管理性に重要なOSPF、BGP、SSHなどのさまざまな制御プロトコルのトラフィックを優先し、同時にスイッチのコントロールプレーン機能に重要ではないタイプのトラフィックを積極的にフィルタリングします。ほとんどのデータトラフィックは、パケット転送例外の結果として CPU に転送される場合、そのようなメカニズムによって厳重にポリシングされます。

一方、ハードウェアレートリミッタとCoPP policing メカニズムは、スイッチのコントロールプレーンの安定性を提供し、常に有効にしておくことを強く推奨します。メカニズムは、データパケットのドロップ、転送遅延、およびネットワーク全体での全体的なアプリケーションパフォーマンスの低下の主な原因の1つである可能性があります。このため、トラフィックフローがネットワークを通過するパスを理解することが重要です。また、ICMPリダイレクト機能を使用できる、または使用することが予想されるネットワーク機器を監視するツールを使用することも重要です。

トラフィックを監視および診断するツール

show ip traffic

Cisco IOSとCisco NX-OSソフトウェアの両方で、CPUによって処理されるトラフィックの統計情報を確認する方法が提供されています。これは次で行われます。 **show ip traffic** コマンドが表示されない場合もあります。このコマンドを使用すると、レイヤ3スイッチまたはルータによるICMPリダイレクトメッセージの受信や生成を確認できます。

```
Nexus7000#show ip traffic | begin ICMP

ICMP Software Processed Traffic Statistics
----------------------------------------
Transmission:
Redirect: 1000, unreachable: 0, echo request: 0, echo reply: 0,

<output omitted for brevity>

ICMP originate Req: 0, Redirects Originate Req: 1000
Originate deny - Resource fail: 0, short ip: 0, icmp: 0, others: 0
Reception:
Redirect: 0, unreachable: 0, echo request: 0, echo reply: 0,

<output omitted for brevity>

Nexus7000#
```

RUN show ip traffic コマンドを数回発行し、ICMPリダイレクトカウンタが増加するかどうかを確認します。

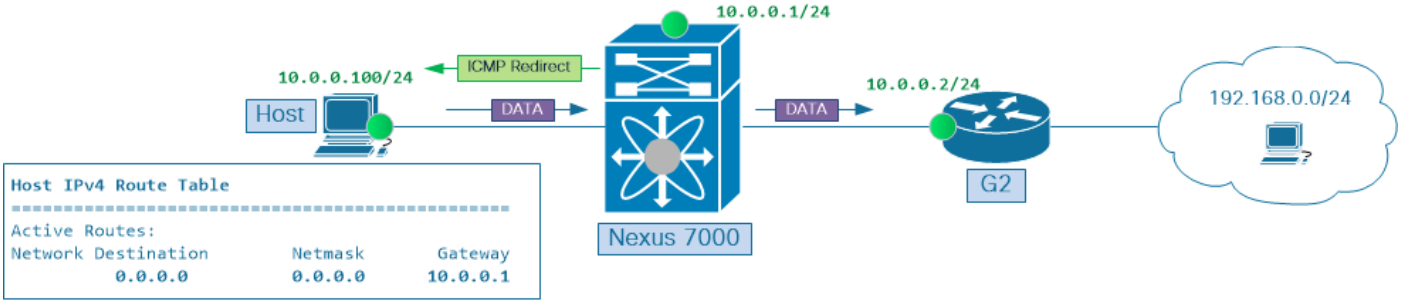
Ethanalyzer

Cisco NX-OSソフトウェアには、トラフィックをキャプチャする組み込みツールがあります flowing ethanalyzerと呼ばれるスイッチCPUとの間で送受信されます。

> **注**：Ethanalyzerの詳細については、『Nexus 7000トラブルシューティングガイドのEthanalyzer』を参照してください。

図6は、図3と同様のシナリオを示しています。ここでは、ネットワークXが192.168.0.0/24ネットワークに置き換えられています。

図6 Ethanalyzerキャプチャの実行



Ethanalyzerキャプチャの実行

ホスト10.0.0.100は、ICMPエコー要求の連続ストリームを宛先IPアドレス192.168.0.1に送信します。ホストは、Nexus 7000スイッチのスイッチ仮想インターフェイス(SVI)10をリモートネットワーク192.168.0.0/24へのネクストホップとして使用します。デモ目的のために、ホストはICMPリダイレクトメッセージを無視するように設定されています。

次のコマンドを使用して、Nexus 7000 CPUによって送受信されるICMPトラフィックをキャプチャします。

```
Nexus7000#ethanalyzer local interface inband capture-filter icmp limit-captured-frames 1000

Capturing on inband
2018-09-15 23:45:40.124077   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.124477     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.124533   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.126344   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.126607     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.126655   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
    2018-09-15 23:45:40.128348   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
    2018-09-15 23:45:40.128611     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
    2018-09-15 23:45:40.128659   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.130362   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.130621     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.130669   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.132392   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.132652     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.132700   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.134347   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.134612     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.134660   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.136347   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.136598     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.136645   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.138351   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request
2018-09-15 23:45:40.138607     10.0.0.1 -> 10.0.0.100   ICMP Redirect (Redirect for host)
2018-09-15 23:45:40.138656   10.0.0.100 -> 192.168.0.1      ICMP Echo (ping) request

...
```

上記の出力のタイムスタンプは、この例で強調表示されている3つのパケット(2018-09-15 23:45:40.128)が同時にキャプチャされたことを示しています。次に、このパケットグループのパケットごとの内訳を示します

- 最初のパケットは入力データパケットであり、この例では ICMP エコー要求です。
  2018-09-15 23:45:40.128348 10.0.0.100 -> 192.168.0.1 ICMP Echo (ping) request

- 2 つ目のパケットは、ゲートウェイによって生成された ICMP リダイレクトパケットです。このパケットはホストに返信されます。

  2018-09-15 23:45:40.128611　10.0.0.1 -> 10.0.0.100　ICMP Redirect (Redirect for host)

- 3 つ目のパケットは、CPU によってルーティングされた後に出力方向でキャプチャされたデータパケットです。前に示していませんが、このパケットのIP TTLは減分され、チェックサムが再計算されます。

  2018-09-15 23:45:40.128659　10.0.0.100 -> 192.168.0.1　ICMP Echo (ping) request

異なるタイプとフローのパケットが多数含まれる大規模なEthanalyzerキャプチャをナビゲートする場合、ICMPリダイレクトメッセージとそれらに対応するデータトラフィックを関連付けることは困難な場合があります。

このような場合には、ICMP リダイレクトメッセージに注目して準最適転送トラフィックフローに関する情報を取得します。ICMPリダイレクト・メッセージには、インターネット・ヘッダーと元のデータグラム・データの最初の64ビットが含まれます。このデータは、データグラムのソースによって、メッセージを適切なプロセスと照合するために使用されます。

Ethanalyzer パケットキャプチャツールで **detail キーワードを使用して、ICMP リダイレクトメッセージの内容を表示し、準最適転送されたデータフローの IP アドレス情報を確認します。**

```
Nexus7000#ethanalyzer local interface inband capture-filter icmp limit-captured-frames 1000
detail

...
Frame 2 (70 bytes on wire, 70 bytes captured)
Arrival Time: Sep 15, 2018 23:54:04.388577000
[Time delta from previous captured frame: 0.000426000 seconds]
[Time delta from previous displayed frame: 0.000426000 seconds]
[Time since reference or first frame: 0.000426000 seconds]
Frame Number: 2
Frame Length: 70 bytes
Capture Length: 70 bytes
[Frame is marked: False]
[Protocols in frame: eth:ip:icmp:ip:icmp:data]
Ethernet II, Src: 00:be:75:f2:9e:bf (00:be:75:f2:9e:bf), Dst: 00:0a:00:0a:00:0a
(00:0a:00:0a:00:0a)
Destination: 00:0a:00:0a:00:0a (00:0a:00:0a:00:0a)
Address: 00:0a:00:0a:00:0a (00:0a:00:0a:00:0a)
.... ...0 .... .... .... .... = IG bit: Individual address (unicast)
.... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
Source: 00:be:75:f2:9e:bf (00:be:75:f2:9e:bf)
Address: 00:be:75:f2:9e:bf (00:be:75:f2:9e:bf)
.... ...0 .... .... .... .... = IG bit: Individual address (unicast)
.... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
Type: IP (0x0800)
Internet Protocol, Src: 10.0.0.1 (10.0.0.1), Dst: 10.0.0.100 (10.0.0.100)
Version: 4
Header length: 20 bytes
Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
0000 00.. = Differentiated Services Codepoint: Default (0x00)
.... ..0. = ECN-Capable Transport (ECT): 0
.... ...0 = ECN-CE: 0
Total Length: 56
Identification: 0xf986 (63878)
Flags: 0x00
0.. = Reserved bit: Not Set
```

```
.0. = Don't fragment: Not Set
..0 = More fragments: Not Set
Fragment offset: 0
Time to live: 255
Protocol: ICMP (0x01)
Header checksum: 0xadd9 [correct]
[Good: True]
[Bad : False]
Source: 10.0.0.1 (10.0.0.1)
Destination: 10.0.0.100 (10.0.0.100)
Internet Control Message Protocol
 Type: 5 (Redirect)
 Code: 1 (Redirect for host)
Checksum: 0xb8e5 [correct]
Gateway address: 10.0.0.2 (10.0.0.2)
Internet Protocol, Src: 10.0.0.100 (10.0.0.100), Dst: 192.168.0.1 (192.168.0.1)
Version: 4
Header length: 20 bytes
Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
0000 00.. = Differentiated Services Codepoint: Default (0x00)
.... ..0. = ECN-Capable Transport (ECT): 0
.... ...0 = ECN-CE: 0
Total Length: 84
Identification: 0xf986 (63878)
Flags: 0x00
0.. = Reserved bit: Not Set
.0. = Don't fragment: Not Set
..0 = More fragments: Not Set
Fragment offset: 0
Time to live: 254
Protocol: ICMP (0x01)
Header checksum: 0xa8ae [correct]
[Good: True]
[Bad : False]
Source: 10.0.0.100 (10.0.0.100)
Destination: 192.168.0.1 (192.168.0.1)
Internet Control Message Protocol
Type: 8 (Echo (ping) request)
Code: 0 ()
Checksum: 0x02f9 [incorrect, should be 0xcae1]
Identifier: 0xa01d
Sequence number: 36096 (0x8d00)

...
```

ICMP リダイレクトのディセーブル化

ネットワーク設計で、スイッチまたはルータに入った同じレイヤ3インターフェイスからトラフィックフローをルーティングする必要がある場合、それに対応するレイヤ3インターフェイスでICMPリダイレクト機能を無効にすると、フローがCPUを経由してルーティングされるのを防ぐことができます。

実際に、ほとんどのネットワークでは、すべてのレイヤ 3 インターフェイス (イーサネット インターフェイスなどの物理的なものとポートチャネル インターフェイスや SVI インターフェイスなどの仮想的なものの両方) で ICMP リダイレクトを事前に無効にすることが推奨されます。 no ip redirects レイヤ3インターフェイスでICMPリダイレクトを無効にするCisco NX-OSインターフェイスレベルコマンド。ICMPリダイレクト機能が無効になっていることを確認するには、次の手順を実行します。

- 保証no ip redirectsコマンドがインターフェイス設定に追加されます。

```
Nexus7000#show run interface vlan 10

interface Vlan10
no shutdown no ip redirects
  ip address 10.0.0.1/24
```

- インターフェイスのICMPリダイレクトのステータスが「disabled」であることを確認します。

```
Nexus7000#show ip interface vlan 10 | include redirects
  IP icmp redirects: disabled
```

- スイッチスーパーバイザから複数のラインカードの1つにインターフェイス設定をプッシュするCisco NX-OSソフトウェアコンポーネントによって、ICMP Redirect enable/disableフラグが0に設定されていることを確認します。

```
Nexus7000#show system internal eltm info interface vlan 10 | i icmp_redirect
  per_pkt_ls_en = 0, icmp_redirect = 0, v4_same_if_check = 0
```

- 1つ以上のラインカードで、特定のレイヤ3インターフェイスのICMPリダイレクト有効/無効フラグが0に設定されていることを確認します。

```
Nexus7000#attach module 7
Attaching to module 7 ...
To exit type 'exit', to abort type '$.'
Last login: Wed Sep 15 23:56:25 UTC 2018 from 127.1.1.1 on pts/0
module-7#

!--- Optionally, jump to non-admin Virtual Device Context (VDC) if verification needs to be done
in one of the custom VDCs
module-7#vdc 6

module-7#show system internal iftmc info interface vlan 10 | include icmp_redirect
icmp_redirect : 0x0 ipv6_redirect : 0x1
```

要約

RFC 792 で規定されている ICMP リダイレクトメカニズムは、マルチポイント ネットワーク セグメントを介した転送パスを最適化するように設計されています。ネットワーク帯域幅が手頃になり、CPUベースのパケットルーティングが専用ハードウェアASICでより高速なレイヤ3パケット転送に進化するにつれて、マルチポイントネットワークセグメントを介した最適なデータ転送の重要性は減少しました。デフォルトでは、ICMPリダイレクト機能はすべてのレイヤ3インターフェイスで有効になっています。スタティックルーティング、ダイナミックルーティング、ポリシーベースルーティングなど、さまざまな転送メカニズムを組み合わせて使用するネットワークでは、ICMPリダイレクト機能を有効のままにして適切に監視しないと、トランジットノードのCPUを使用して実稼働トラフィックを処理する必要がなくなります。これにより、実稼働トラフィックフローとネットワークインフラストラクチャのコントロールプレーンの安定性の両方に大きな影響が生じる可能性があります。

ほとんどのネットワークでは、ネットワーク インフラストラクチャのすべてのレイヤ3インターフェイスで ICMP リダイレクト機能を事前に無効にすることが推奨されます。これは、マルチポ

イントネットワークセグメントを通じてより優れた転送パスが存在する場合に、レイヤ3スイッチおよびルータのCPUで処理される実稼働データトラフィックのシナリオを回避するのに役立ちます。

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。