WCCP を使用したパス MTU ディスカバリの WSA の動作

内容

概要

このドキュメントでは、設定に Web Cache Communication Protocol（WCCP）とパス最大伝送ユニット（MTU）ディスカバリの両方が含まれる場合に、ルータがパケットをドロップする問題について説明し、その問題に対する解決策を提供します。

背景説明

事前フェーズ

個別に見ると、特定の問題の処理に優れた機能は数多くあります。ただし、2 つまたは 3 つの機能を組み合わせると、動作が不安定になることがあり、それを正常に動作させるために別の機能または回避策を導入する必要に迫られます。たとえば、スパニング ツリーと併せて Open Shortest Path First（OSPF）とレイヤ 2（L2）コンバージェンスを使用すると、OSPF よりも処理時間が長く（20 秒）なりますが（最小限の dead 間隔を使用する場合は 1 秒）、スパニング ツリーをマルチ スパニング ツリー（MST）に置き換えると、元のように正常に機能します。

WCCP と Path MTU Discovery の間でも、同じような相互運用性の動作が観察されています。Generic Routing Encapsulation（GRE）のヘッダーの問題であると多くの人が考えています。しかしながら、このドキュメントでは真の原因について説明します。

Path MTU Discovery と WCCP の異なる動作方式

Path MTU Discovery

各ラインには、通過できるパケットの大きさに制限があります。サポートされているよりも大きなパケットを送信すると、そのパケットはドロップされます。途中に存在する L3 デバイス (ルータ) は、その役割の 1 つとして、エンドツーエンドの通信が各ラインの能力に対して透過的になるように、あるラインから他のラインへ送信される大きなパケットを処理し、適当な大きさに細分化することです。

ただし、パケットを細分化できないようにエンド ホストが設定されている場合があります (たとえば、暗号化ファイル、音声通話など)。 この情報は、IP ヘッダー内の Don't Fragment (DF) ビットを使用して伝達されます。ルータはこのようなパケットをドロップしますが、Internet Control Message Protocol (ICMP) メッセージ (type 3-Destination unreachable、code 4 - fragmentation needed、but DF bit set) を使用してエンド ホストに報告しようとします。 このようしてホストは、将来、より小さなパケットを送信する必要があることを認識します。

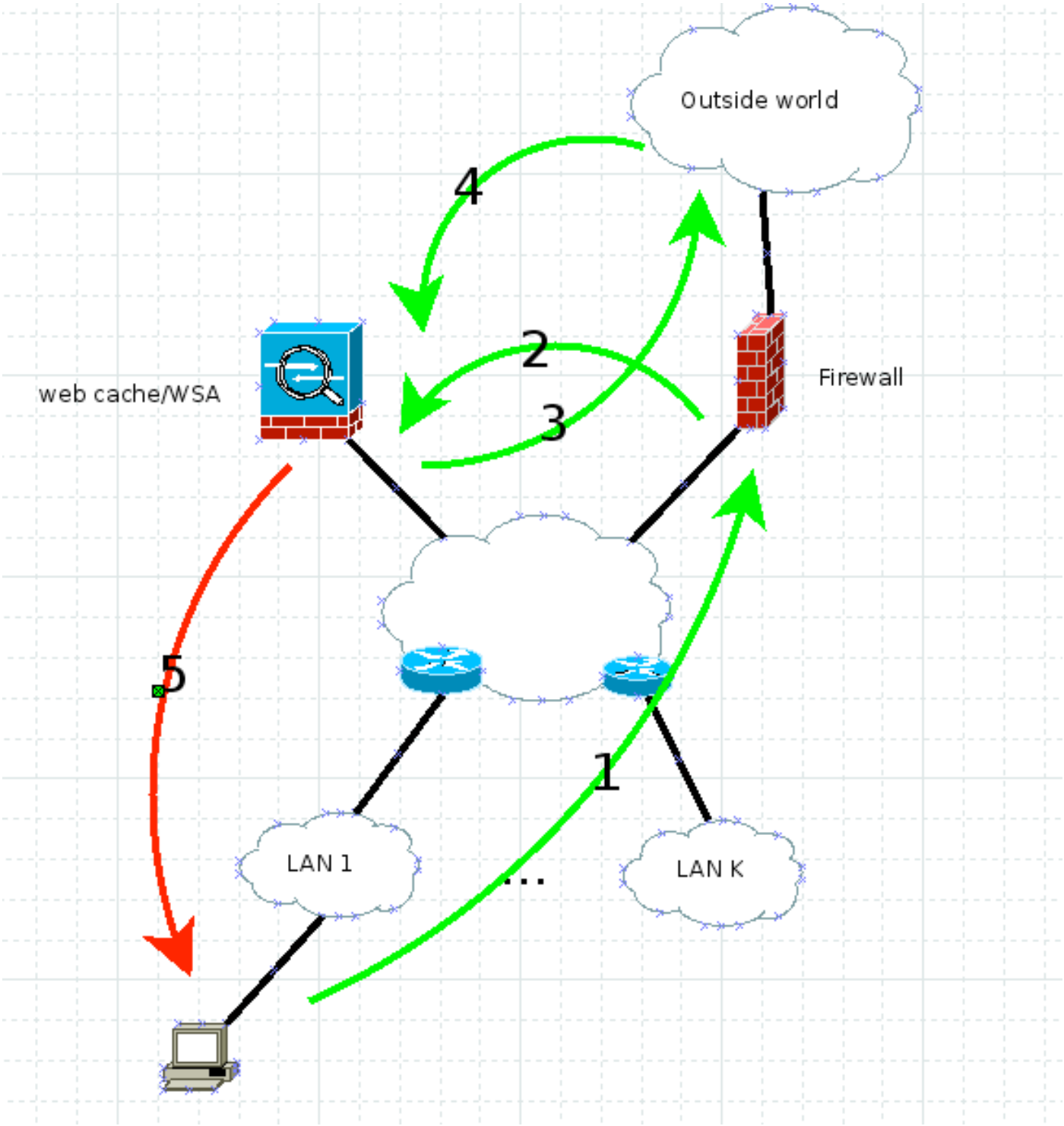
これが、Path MTU Discovery の核心的な仕組みです。パケットが送信先まで届くか、前述のように ICMP の報告を受け取るかどうかを判断するために、DF ビットを設定した大規模なパケットを送信できます。送信可能な最大パケット サイズが決まったら、その後の通信にはそのサイズが使用されます。詳細については、RFC 1191 を参照してください。

Web セキュリティ アプライアンス (WSA) は、デフォルトで Path MTU Discovery を使用します。そうすることで、生成されたすべてのパケットに、デフォルト設定で DF ビットが付加されます。

WCCP

トラフィックに関する知識がなくとも、Web トラフィックに対するセキュリティをネットワークに導入する必要がある場合は、見えないプロキシ経由でトラフィックの通信を行うようにします。WCCP とは、傍受するデバイス (ルータ/ファイアウォール) と Web キャッシュエンジン/プロキシ間の通信に使用されるプロトコルで、ここでは WSA です。

この図は、トラフィックがこのシナリオでどのように流れるかを示しています。



次のように機能します。

1. クライアントが、IP 送信元、IP アドレス (クライアント IP アドレス) 、および送信先サーバの IP アドレスを指定して HTTP GET を送信します。

2. ファイアウォールやルータが HTTP GET を傍受し、WCCP GRE または純粋な L2 を経由して Web キャッシュ/WSA に転送します。送信元は引き続きクライアントの IP アドレスで、宛先は Web サーバの IP アドレスです。

3. WSA がリクエストを検査し、それが正当である場合は、Web サーバ向けにそれをミラーリングします。ここで、宛先 IP アドレスは Web サーバの IP アドレスで、送信元 IP アドレスは、クライアント IP アドレスのスプーフィングを有効にしたかどうかに基づいて、WSA またはクライアントのいずれかです。この例では、いずれの場合もリターン トラフィックは

WSA に達する必要があるためにそのいずれかは問題ではありません。

4. リターン トラフィックは WSA で検査されます。

5. WSA は、送信元 IP アドレス、常に Web サーバの IP アドレス (つまり、クライアントから疑われない)、および宛先クライアント IP アドレスを使用してクライアントに応答を送信します。

問題

図のルータのいずれかがトラフィックを断片化する必要がある場合はどうなりますか。WSA は、パケット番号 5 に DF ビットを配置しますが、これには断片化が必要です。ルータはそれをドロップし、断片化が必要だが DF ビットが設定されている (ICMP タイプ 3 コード 4) ことを送信者に通知します。 結局のところ、ここでは RFC 1191 が動作する必要があり、送信側はそのパケット サイズを小さくする必要があります。

WCCP では、送信元 IP アドレスは Web サーバの IP アドレスなので、この ICMP が WSA に送信されることはありません。むしろ、実際の Web サーバに到達しようとします (この一番下のルータは WCCP を認識しません)。 これが、場合によっては、WCCP および Path MTU Discovery が組み合わさって、ネットワーク設計を壊してしまう仕組みです。

解決方法

この問題を解決するには、次の 4 つの方法があります。

- 実際の MTU を発見してから、WSA で **etherconfig を使用してインターフェイスの MTU を小さくします。**TCP ヘッダーは 60、IP は 20 で、ICMP を使用する場合は IP ヘッダーに 8 バイトを追加することを忘れないでください。

- Path MTU Discovery を無効にします (**pathmtudiscovery CLI WSA コマンド) 。** これにより、536 の TCP MSS が発生し、パフォーマンス上の問題が起こる可能性があります。

- NWSA とクライアントの間に L3 の断片化が存在しないようにネットワークを変更します。

- **ip tcp mss-adjust 1360 (またはその他の計算された番号) コマンドを、関連するインターフェイスまでの途中にある各 Cisco ルータで実行します。**

追加情報

この問題はまだ調査中ですが、クライアントの中に明示的にプロキシを数分間設定してからそれを取り除くと、以後 4 ~ 5 時間は問題が解決することがわかりました。これは、明示的なモードでは、NWSA とクライアントとの間で Path MTU Discovery の仕組みが機能することが原因です。WSA がパス MTU を発見すると、発見した TCP MSS とともに、そのパス MTU を参照用に内部テーブルに格納します。このテーブルは 4 ~ 5 時間ごとに更新されるようなので、それだけの時間が経過すると、この解決策は機能しなくなります。