在SD-WAN中為SDA部署配置最佳ISE IP MTU

目錄

簡介

必要條件

- 需求
- 採用元件:

背景資訊

問題描述

- 示例拓撲

練習1:MTU缺口 — SDA邊界至SD-WAN邊緣

挑戰1的解決方案：

練習2:MTU壓縮 — SD-WAN重疊中的ISE流量

- 封包結構和封裝額外負荷：

挑戰2的解決方案：主動ISE IP MTU配置

- ISE配置（通過CLI的示例）：

結論

- 標準和參考

簡介

本文檔介紹在使用SD-WAN連線SDA站點時，最大傳輸單元(MTU)問題如何影響SDA中的微分段。

必要條件

需求

思科建議您瞭解以下主題：

- 思科軟體定義存取(SDA)
- 思科軟體定義廣域網路(SD-WAN)
- 思科身分識別服務引擎(ISE)

採用元件:

本文檔中的資訊基於SDA、SDWAN和ISE。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

背景資訊

現代企業網路越來越多地利用SDA實現細粒度細分和一致的策略實施。為了連線分散式SDA站點，通常使用Cisco SD-WAN，通過多種底層網路提供靈活、安全和最佳化的傳輸。作為此架構的核心，ISE提供關鍵身份驗證、授權和記帳(AAA)服務，以及動態策略分配(例如，安全組標籤(SGT)和可下載ACL)。

雖然功能強大，但是這些強大技術的整合可能會帶來微妙但卻有效的配置挑戰。關鍵網路切換點和跨SD-WAN重疊的MTU處理是此類問題的主要方面。本文討論了兩種可能中斷網路運行的常見MTU不匹配情況：

1. SDA邊界節點和SD-WAN邊緣裝置之間的MTU間隙。
2. 通過SD-WAN重疊的ISE發起流量的MTU限制。

正確的MTU對齊對於防止資料包分段問題或靜默丟棄至關重要，可確保可靠的身份驗證、策略實施和整體網路穩定性。如果無法解決這些問題，可能會導致間歇性的連線和策略實施故障令人困惑，從而耗費大量的故障排除工作。

MTU未對齊的常見症狀

不一致的MTU可能以多種方式表現出來，通常會導致難以診斷的問題：

- 間歇性RADIUS身份驗證失敗或超時：尤其值得注意的是生成較大RADIUS資料包的策略（例如，具有大量AV配對或證書的策略）。
- 端點無法接收或應用可下載ACL(dACL)或TrustSec策略(SGT/SGACL):這些原則通常在大型RADIUS封包中傳達。
- 通過身份驗證的客戶端建立會話速度緩慢：由於應用層重新傳輸。
- RADIUS重新傳輸過多：可在ISE日誌或網路訪問裝置(NAD)上觀察到。
- 策略傳播不一致：在ISE中進行的策略更改可能不會始終傳播到遠端SDA站點中的所有NAD。
- 資料包捕獲差異：捕獲可以顯示ISE傳送大型資料包(例如>1450位元組)，並設定Do Not Fragment(DF)位元，但是沒有相應的響應或NAD或SD-WAN思科邊緣路由器的ICMP「需要分段」錯誤。
- 遞增資料包丟棄計數器：在資料中心(DC)思科邊緣路由器的入口介面上觀察到來自ISE且目的地為SDA站點的流量，或在面向SDA邊界的SD-WAN思科邊緣路由器介面上觀察到反向流量。

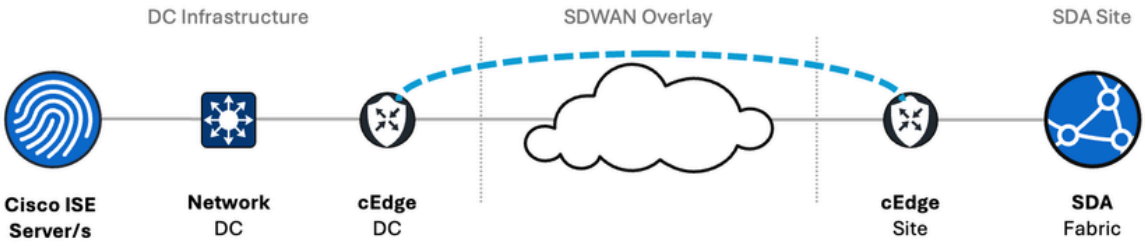
問題描述

典型的企業部署
考慮常見的企業拓撲：

- Cisco ISE伺服器：部署在與DC網路基礎設施連線的集中式資料中心(DC)或區域中心。
- DC基礎設施：包括ISE伺服器所連線的DC核心或匯聚交換機。

- SD-WAN覆蓋：DC思科邊緣路由器路由器通過底層傳輸網路(例如，互聯網、MPLS)建立到遠端SDA站點的Cisco邊緣路由器路由器的SD-WAN隧道(通常為IPsec)。
- SDA站點：遠端站點思科邊緣路由器路由器連線到本地SDA交換矩陣，包括交換矩陣邊緣節點、邊界節點、無線LAN控制器(WLC)，並最終連線到終端。

示例拓撲



練習1:MTU缺口 — SDA邊界至SD-WAN邊緣

思科SDA設計原則（通常通過LAN自動化實施）促進了所有交換矩陣裝置上的9100位元組（巨型幀）的園區範圍的MTU。這包括Catalyst 9000系列邊界節點，並確保乙太網巨型幀在交換矩陣中有效地傳輸。因此，SDA邊界節點上的第3層或SVI切換介面預設為此更大的MTU。

相反，SD-WAN邊緣裝置（如Catalyst 8000系列）通常預設為1500位元組的介面MTU。這是連線到外部網路(如Internet服務提供商(ISP))的介面的標準配置，在這些網路中，巨型幀支援不常見或未啟用。

這種差異直接導致了潛在的失敗：嘗試將大於1500位元組的IP資料包傳送到其接收介面配置為1500位元組MTU的SD-WAN邊緣的SDA邊界。

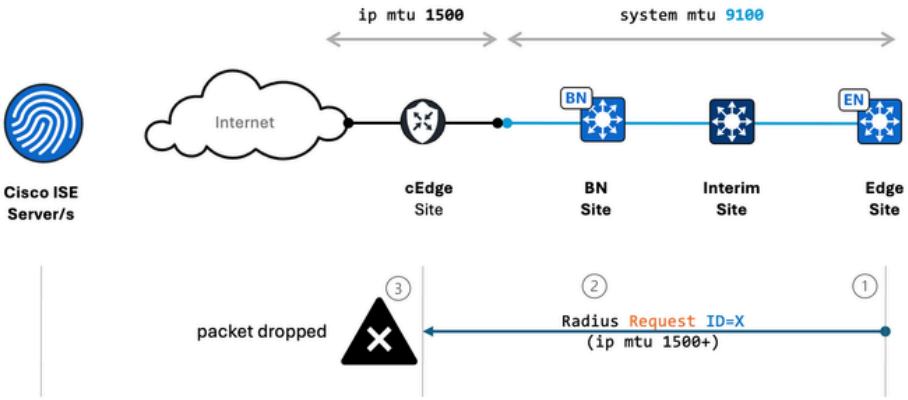
此型別的MTU不匹配是SDA部署中的常見缺陷，在配置過程中通常很容易被忽略。更具有挑戰性的是，與如何在執行Cisco IOS-XE®的Catalyst 9000交換器上產生RADIUS要求相關的某些行為，只能導致在特定及緊急情況下才會暴露這些問題。

例如，在由會話管理器守護程式(SMD)進程處理的終端使用者身份驗證過程中生成的RADIUS請求被硬編碼為以1396位元組對資料包進行分段。相反，檢索TrustSec策略(例如安全組訪問控制清單(SGACL))時涉及的RADIUS請求是由思科互聯作業系統守護程式(IOSd)子元件生成的。這些資料包可感知MTU，並且可避免對資料包進行分段，除非其大小超過系統MTU（通常高達9100位元組）。

因此，與MTU不相符相關的問題僅在使用Cisco TrustSec(CTS)下載策略時才會變得明顯。此外，在使用者身份驗證期間，SDA邊緣裝置下載的基於角色的訪問控制清單(RBACL)集可能會有所不同，具體取決於其他標籤已經存在哪些SGACL策略。實際上，交換機只下載策略集的非重疊部分。

綜合起來，這些行為可能會導致不可預測和不一致的結果，從無聲故障到不完整策略下載，具體取

決於SGACL策略的規模、當前系統條件，並最終取決於路徑上的MTU失調。



SDA邊界通過SD-WAN邊緣向ISE轉發大型RADIUS資料包(例如,1600位元組)，出現以下情況：

1. SDA邊界及其9100 MTU介面會傳送1600位元組的IP封包。
2. SD-WAN思科邊緣路由器在其1500 MTU介面上收到此資料包。
3. 但是，如果沒有在這些RADIUS封包上設定「不分段(DF)」位元，SD-WAN思科邊緣路由器通常會在輸入時將其捨棄，只是因為它們相對於其已設定的介面MTU「過大」。 它不會進入IP轉送邏輯的階段，在此階段它可以考慮將其分段（如果DF位元允許）。

這種靜默丟棄會導致嚴重的故障排除問題，特別是因為問題是定向的（SDA到SD-WAN/ISE）。

類似的MTU不匹配可能發生在資料中心(DC)核心或枝葉交換機，這些交換機通常配置為支援巨型幀(例如，MTU 9000+)，以提高內部DC流量效率。但是，如果流量被傳遞給配置了標準MTU(例如1500位元組)的SD-WAN DC思科邊緣路由器面向LAN的介面，這種不匹配可能會導致分段或丟包，尤其是對於從DC網路流入SD-WAN交換矩陣的流量。

挑戰1的解決方案：

將SDA邊界切換介面（物理或SVI）的IP MTU與對等的SD-WAN Cisco Edge路由器介面對齊，通常為1500位元組。

配置示例（在SDA邊界節點上）：

<#root>

```
!
interface Vlan3000  // Or your physical handoff interface, for example, TenGigabitEthernet1/0/1
 description Link to SD-WAN cEdge Router
 ip address 192.168.100.1 255.255.255.252


ip mtu 1500

  // Align with SD-WAN cEdge receiving interface MTU
!
```

重要考慮因素：Catalyst 9000邊界上的分段

Catalyst 9000系列交換器作為SDA邊界節點，支援硬體資料平面中原生IP封包的IP分段。將切換介面上的ip mtu降低到1500 不會導致效能降低，因為來自或穿過需要它的邊界的資料流的基於軟體的分段。交換器在進入此特定介面之前，高效地對大於1500位元組的IP封包進行分段（如果DF位元已清除），而不會傳送到CPU。

但是，必須注意的是，Catalyst 9000交換器通常不支援VXLAN封裝流量的分段。此限制對重疊流量至關重要，但不會影響所述的RADIUS身份驗證場景，因為SDA邊界和外部ISE之間的RADIUS通訊通常發生在底層（本地IP路由）中。（VXLAN重疊的MTU注意事項是一個單獨且複雜的主題，詳細內容見相關思科SDA設計手冊）。

SDA邊界的主動式MTU與SD-WAN Cisco邊緣路由器切換至關重要。

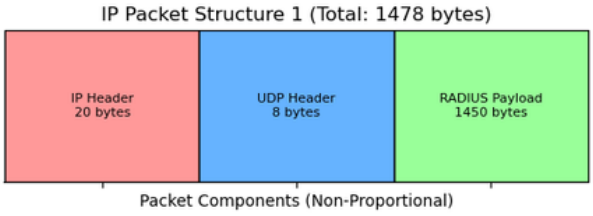
練習2:MTU壓縮 — SD-WAN重疊中的ISE流量

即使單個物理介面(例如ISE網路介面卡(NIC)、交換機埠或路由器介面)設定為標準1500位元組IP MTU，SD-WAN重疊本身也會引入封裝開銷。此額外負荷會消耗1500位元組限制的一部分，從而降低原始IP封包的有效MTU(從ISE的角度來看是「負載」)。

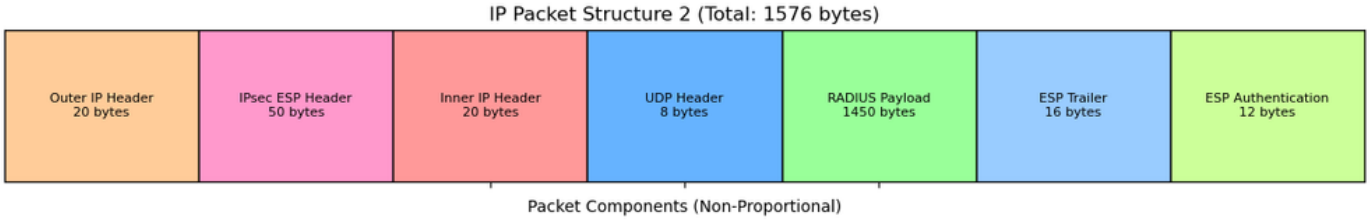
封包結構和封裝額外負荷：

當從ISE伺服器(例如，RADIUS訪問接受資料包)傳送至SDA站點中的網路訪問裝置(NAD)時，它會穿越SD-WAN重疊並封裝。常見的封裝堆疊在通道模式下涉及IPsec，可能是通過UDP進行NAT穿越(NAT-T)。

- 來自ISE的原始資料包（內部資料包）：
  例如，具有1450位元組負載的RADIUS封包+ 8B UDP + 20B內部IP = 1478位元組。

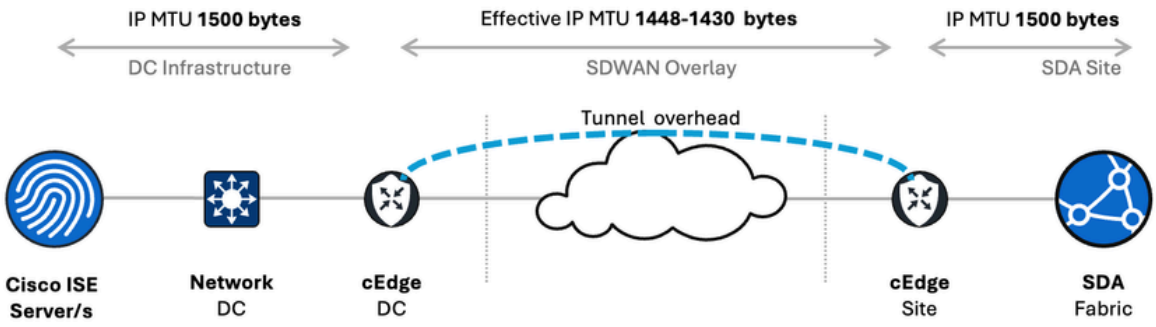


- 考慮在隧道模式下使用IPsec ESP，可能使用用於NAT-T的UDP封裝：



- 總開銷可能因特定IPsec密碼、驗證機制和其他重疊功能(如GRE（如果使用）而異。典型計算：
  - 外部IP標頭(IPv4):20 位元組

- UDP報頭（如果用於NAT-T的ESP over UDP）：8 位元組
- ESP報頭：~8位元組
- ESP IV(例如，適用於AES-CBC):~16位元組（如果適用）
- ESP身份驗證(例如，截斷HMAC-SHA256):~12-16位元組
- 常見IPsec估計開銷：~52-70位元組（可能更高，若有所有選項，最高可達~80位元組或更多）。



如果實體連結MTU為1500位元組，則來自ISE的原始IP封包的可用負載MTU變成：1500位元組 — SD-WAN額外負荷。
例如，1500 - 70 = 1430位元組。

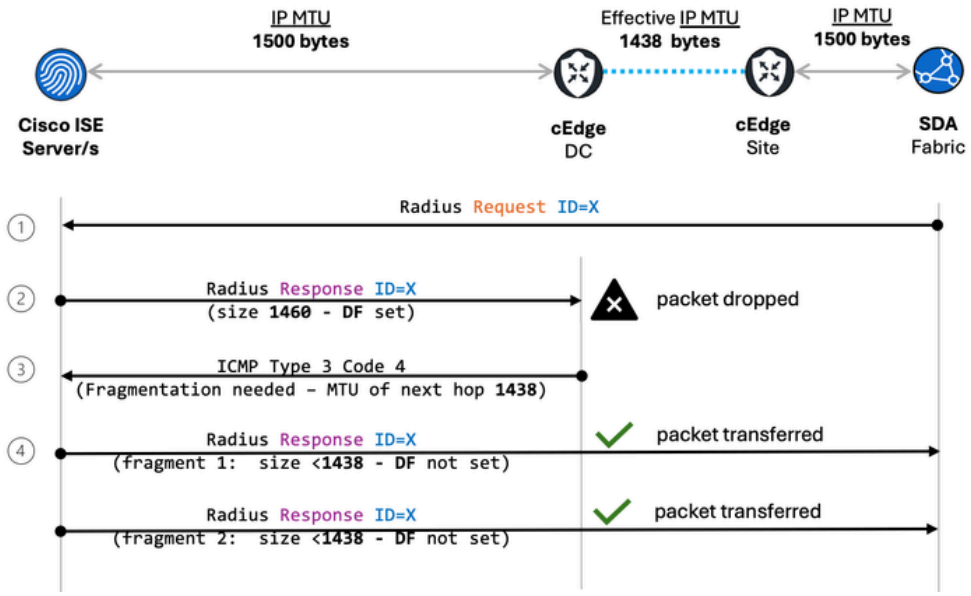
封包超過有效MTU時的行為：

1. ISE建立封包（DF位元異常）：
   - 預設情況下，ISE裝置的基礎Linux作業系統會為其源的所有小於或等於其配置的介面IP MTU(例如，1500位元組)的資料包設定Do Not Fragment(DF)標簽。
   - 此DF位元的用途：ISE（通過其OS）主動設定DF位元，主要是利用路徑MTU發現(PMTUD)流程，稍後將對此進行介紹。這麼做可讓ISE動態得知目的地的實際PMTU（若目的地的實際PMTU小於自己的介面MTU）。
   - 大於介面MTU的資料包行為：如果ISE需要傳送的IP資料包大於其配置的介面IP MTU，則行為取決於其Linux作業系統。通常，OS會在傳輸之前對封包進行分段，並清除這些產生的片段上的DF位元（設定DF=0）。此分段是OS級別的函式，不直接由ISE應用程式碼本身驅動。
   - 與網路裝置的關鍵區別：ISE的此預設行為（設定DF=1，即使對於適合其介面MTU中的非分段資料包）與許多傳統網路裝置（路由器、交換機）有很大不同。 網路裝置通常不會在其發起或轉發的資料包上設定DF位元，除非明確配置為這樣做，或者如果被轉發的資料包已設定DF位元，或者為特定協定設定DF位元。如果封包超過下一個躍點的MTU（和DF=0），則通常會在預設情況下允許分段。
   - 故障排除複雜性：這種不對稱性(ISE到NAD流量通常預設情況下為DF=1，而NAD到ISE流量可以為DF=0（除非該NAD設定其原因） — 在故障排除期間會引入額外的複雜性層。工程師可以觀察到不同的分段行為和PMTUD互動，具體取決於通訊流的方向。
2. 資料包到達入口思科邊緣路由器(DC):DC思科邊緣路由器從ISE接收IP資料包。
3. Cisco Edge路由器的封裝和MTU檢查：Cisco Edge路由器嘗試為SD-WAN隧道封裝資料包。

- 如果原始資料包大小為SD-WAN封裝開銷超過思科邊緣路由器的出站物理介面MTU(例如，1500位元組)，並且在來自ISE的原始（內部）資料包上設定DF位元，則思科邊緣路由器不得對內部資料包進行分段。
- 思科邊緣路由器應捨棄封包。
- 關鍵是，思科邊緣路由器還應將ICMP「無法到達目的地 — 需要分段和設定DF位元」（型別3，代碼4）訊息傳送回來源(ISE)，指出下一個躍點的MTU（通道的有效MTU）。

4. 路徑MTU探索(PMTUD)程式：收到此ICMP「需要分段」訊息時，ISE（來源作業系統）應降低該特定目的地路徑的PMTU估計值。它會快取此資訊，並以適合新發現的PMTU的較小封包重新傳送資料。

PMTUD流程圖：



其中PMTUD通訊中斷：
PMTUD在理論上是穩健的，但在實踐中可能會失敗：

- ICMP過濾：中間防火牆或安全策略通常會阻止ICMP消息，阻止「需要分段」消息到達ISE。
- 思科邊緣路由器上的控制階段管制(CoPP):Cisco Edge Router路由器使用CoPP保護其CPU。生成ICMP錯誤消息是一項控制平面任務。在負載繁重或存在許多超大資料包的情況下，CoPP可以限制速率或丟棄ICMP生成。ISE永遠不會收到反饋。
- 靜默丟棄：如果ISE沒有收到ICMP「需要分段」消息，它仍然不知道路徑限制。它繼續傳送已設定DF位元的大型資料包，導致入口思科邊緣路由器以靜默方式丟棄這些資料包。這會導致應用層超時和重新傳輸(例如RADIUS)。
- 對ISE服務的影響：大型RADIUS Access-Accept封包（承載dACL、廣泛的AVP、SGT資訊）尤其容易受影響。表現形式包括：
  - 間歇或完全身份驗證失敗。
  - 未收到正確網路訪問策略或SGT的終端。

- ISE和NAD之間的策略同步不完整或失敗。

挑戰2的解決方案：主動ISE IP MTU配置

鑑於PMTUD不可靠，主動式方法最適合於關鍵服務（如ISE）。將ISE網路介面上的IP MTU配置為可安全容納最大預期SD-WAN重疊開銷的值。這可確保ISE不會發起本身過大而不需要中間裝置（如果DF=1則禁止）通過SD-WAN重疊的IP資料包（設定DF位）。

計算和設定建議的ISE IP MTU:

1. 建立基本物理MTU:對於路徑沿線的標準乙太網介面，這通常為1500位元組。
2. 確定最大SD-WAN封裝開銷：
   - 準確計算或保守估計特定SD-WAN重疊（IPsec、GRE、VXLAN、MPLSoGRE等）引入的總開銷。 有關您選擇的協定和選項的精確數字，請參閱供應商文檔。

元件	開銷示例（位元組）	備註
基本物理MTU	1500	物理鏈路上的標準乙太網
減：SD-WAN開銷		
外部IP標頭(IPv4)	20	
UDP報頭（用於NAT-T）	8	如果ESP封裝在UDP中
ESP報頭	~8-12	
ESP IV(例如AES-CBC)	~16	因加密演算法而異
ESP驗證(例如SHA256)	~12-16	隨身份驗證演算法而變化（例如，某些演算法為96位）
其他重疊（GRE等）	變數	新增SD-WAN封裝堆疊的一部分
總估計開銷	~68 - 80+位元組	部署的所有相關元件的總和
有效路徑MTU	~1432 - 1420位元組	基本物理MTU — 估計的總開銷

3. 建議的ISE IP MTU配置：
   - 採用計算得到的有效路徑MTU（例如，示例中的1420位元組）。
   - 減去額外的安全餘量（例如20-70位元組），以說明次要的未入帳第2層標頭，或提供緩衝區。
   - Cisco SD-WAN等解決方案可以為每個站點到站點隧道單獨執行路徑MTU(PMTU)發現。此機制每20分鐘自動運行一次，以測試隧道的IP MTU，並根據每個站點的當前傳輸條件動態調整其IP MTU。因此，站點之間的MTU值可能不同，並且可能隨時間而更改。
   - 在此類情況下，ISE介面通常安全且建議的IP MTU介於1350和1400位元組之間

1350位元組的IP MTU通常是一個非常穩健的起始點

ISE配置（通過CLI的示例）：

此命令在思科ISE裝置CLI上為每個相關網路介面執行。

<#root>

```
!
interface GigabitEthernet0  ! Or the specific interface used for RADIUS/SDA communication
```

```
 ip mtu 1350
!
```

ISE IP MTU更改的重要操作注意事項：

- 需要服務重新啟動：對ISE介面應用ip mtu命令後，將提示使用者重新啟動ISE應用服務。這是一個影響服務的更改，必須在計畫的維護時段內進行計畫。有關操作詳細資訊，請參閱官方的Cisco ISE文檔。
- 應用於所有ISE節點：此IP MTU調整必須一致地應用於部署中通過SD-WAN與NAD通訊的所有ISE節點(主PAN、輔助PAN、策略服務節點(PSN))。MTU設定不一致會導致不可預測的行為。
- 全面測試：在生產中實施之前，在實驗室或試運行部署中嚴格測試此更改。使用具有各種封包大小和DF位元設定的ping等工具來驗證端對端MTU處理：
    - 基於Linux的系統：

```
ping
        -s
         -M do
```

(附註：-s指定ICMP負載大小。總IP封包大小=負載+ 8B ICMP Hdr + 20B IPv4的IP Hdr)

    - Windows:

```
ping
        -f -l
```

(附註：-l指定ICMP負載大小。)

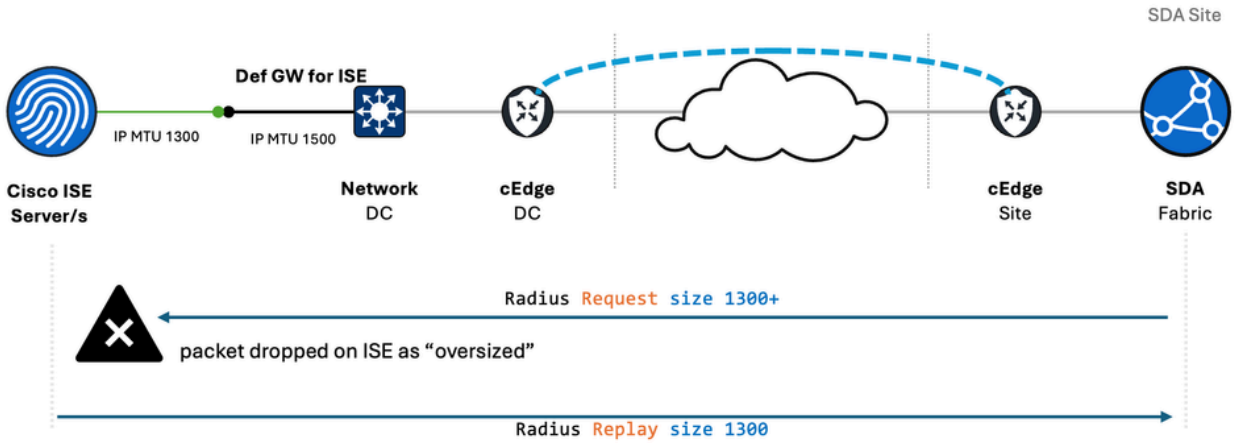
- Cisco IOS/Cisco IOS-XE®

```
ping

        size

         df-bit
```

- ISE第一個路由點 — 調整ISE介面上的IP MTU值時，請確保資料中心中的第一個路由點（尤其是與ISE子網關聯的第3層介面）也配置相同的IP MTU值。
  這有助於防止類似練習1中所述的情況，其中MTU不匹配導致ISE將傳入資料包視為過大並丟棄它們。
  例如，如果ISE介面的MTU減少(例如，1300)，但第一個路由點仍使用預設MTU 1500進行配置，則傳送到ISE大於1300位元組但小於1500位元組的資料包不會分段，並由ISE丟棄，如練習1所示。
  此外，請確保第一個路由點在需要時能夠執行分段，並且這樣做不會導致效能下降。

- 更新整個傳輸路徑及兩個方向的MTU — 更新ISE上的IP MTU設定時，必須考慮整個傳輸路徑及兩個方向的MTU。如果ISE上配置的MTU值與第一跳網關的第3層介面的MTU不一致，可能會出現類似的問題，如挑戰#1中所述。

  例如，如果ISE MTU降低到1300位元組，而預設1500位元組MTU仍配置在預設網關上，則網路裝置通常生成的1300到1500位元組之間的資料包可能會被ISE丟棄為過大。



要避免此問題，請始終確保ISE上的MTU更改映象到第一跳網關上，最好反映在同一第3層網段內的所有終端主機上。這有助於保持端到端MTU一致性，並防止意外丟包。

結論

按照SD-WAN封裝和SDA邊界的MTU調整所施加的有效傳輸層MTU限制來調整Cisco ISE伺服器上的IP MTU設定到SD-WAN Cisco Edge Router切換不僅是一項建議，而且是在現代分段企業網路中確保AAA服務的穩定性、可靠性和效能的關鍵先決條件。雖然路徑MTU探索是一種重要機制，但在SD-WAN環境中，它的實際有效性可能會受到ICMP過濾或控制平面策略等因素的影響。

通過主動在ISE上配置降低的IP MTU(例如，1350-1400位元組)，網路架構師和工程師可以顯著降低與MTU相關的資料包丟棄風險，從而實現更可預知和更恢復的網路操作。這一點在思科SDA部署中尤其重要，ISE在這些部署中協調複雜的微分段和動態策略實施，而後者通常依賴於成功交付潛在的大型控制平面消息。跨所有ISE節點進行勤奮規劃、全面測試和一致的配置是成功且無故障的部署的關鍵所在。

標準和參考

要更深入地瞭解，請參閱官方標準和思科文檔：

RFC:

- RFC 1191:路徑MTU探索
- RFC 791:Internet協定(IP) — 定義IP報頭，包括不分段(DF)位。
- RFC 8200:IPv6規範（如果使用IPv6則相關，包括類似的PMTUD概念）。
- RFC 4459:網內通道(VPN)的MTU和分段問題 — 直接解決VPN環境中的常見MTU問題。

思科檔案：

- 思科SDA設計和部署指南：有關交換矩陣MTU建議和邊界節點配置的資訊。
- Cisco SD-WAN設計和配置指南：有關封裝額外負荷、通道介面MTU和SD-WAN光纖中的PMTUD注意事項的詳細資訊。
- Cisco Catalyst 9000系列交換機配置指南：獲取有關MTU設定和分段功能的平台特定詳細資訊。
- 思科身份服務引擎(ISE)管理員和CLI指南：有關介面配置的資訊，包括ip mtu命令和服務重新啟動含義。

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。