

在SD-WAN中为SDA部署配置最佳ISE IP MTU

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件:](#)

[背景信息](#)

[问题说明](#)

[示例拓扑](#)

[练习1:MTU缺口 — SDA边界到SD-WAN边缘](#)

[挑战1的解决方案:](#)

[练习2:TINC Garbage Ltd.通过SD-WAN重叠的MTU Squeeze - ISE流量](#)

[数据包结构和封装开销:](#)

[挑战2的解决方案:主动ISE IP MTU配置](#)

[ISE配置 \(通过CLI的示例\):](#)

[结论](#)

[标准和参考](#)

简介

本文档介绍在使用SD-WAN连接SDA站点时，最大传输单元(MTU)问题如何影响SDA中的微分段。

先决条件

要求

Cisco 建议您了解以下主题：

- 思科软件定义访问(SDA)
- 思科软件定义广域网(SD-WAN)
- 思科身份服务引擎(ISE)

使用的组件:

本文档中的信息基于SDA、SDWAN和ISE。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

现代企业网络越来越多地利用SDA进行精细微分段和一致的策略实施。要连接分布式SDA站点，通常使用思科SD-WAN，通过各种底层网络提供灵活、安全和优化的传输。作为此架构的核心，ISE提供关键身份验证、授权和记帐(AAA)服务以及动态策略分配(例如，安全组标记(SGT)和可下载ACL)。

虽然功能强大，但是这些强大技术的集成可能会带来微妙而又有效的配置挑战。在关键网络切换点和跨SD-WAN重叠的MTU处理是解决此类问题的主要方面。本文讨论两种常见的MTU不匹配情况，它们可能会中断网络运行：

1. SDA边界节点和SD-WAN边缘设备之间的MTU间隙。
2. 通过SD-WAN重叠的ISE发起流量的MTU限制。

正确的MTU对齐对于防止数据包分段问题或静默丢弃、确保可靠的身份验证、策略实施和整体网络稳定性至关重要。如果无法解决这些问题，可能会导致间歇性的连接和策略实施故障令人困惑，从而耗费大量的故障排除工作。

MTU未对齐的常见症状

不一致的MTU可能以各种方式表现出来，通常会导致难以诊断的问题：

- 间歇性RADIUS身份验证失败或超时：对于生成较大的RADIUS数据包的策略（例如，具有大量AV对或证书的策略）尤其明显。
- 无法接收或应用可下载ACL(dACL)或TrustSec策略(SGT/SGACL)的终端：这些策略通常在大型RADIUS数据包中传输。
- 身份验证客户端的会话建立缓慢：由于应用层重新传输。
- 过多的RADIUS重新传输：可在ISE日志或网络访问设备(NAD)上观察到。
- 策略传播不一致：在ISE中进行的策略更改可能不会始终传播到远程SDA站点中的所有NAD。
- 数据包捕获差异：捕获可以显示ISE发送设置了不分段(DF)位的大数据包(例如，>1450字节)，但是没有来自NAD或SD-WAN思科边缘路由器的相应响应或ICMP“需要分段”错误。
- 递增数据包丢弃计数器：在数据中心(DC)思科边缘路由器的入口接口上观察到来自发往SDA站点的ISE的流量，或在面向SDA边界的SD-WAN思科边缘路由器接口上观察到反向方向的流量。

问题说明

典型的企业部署

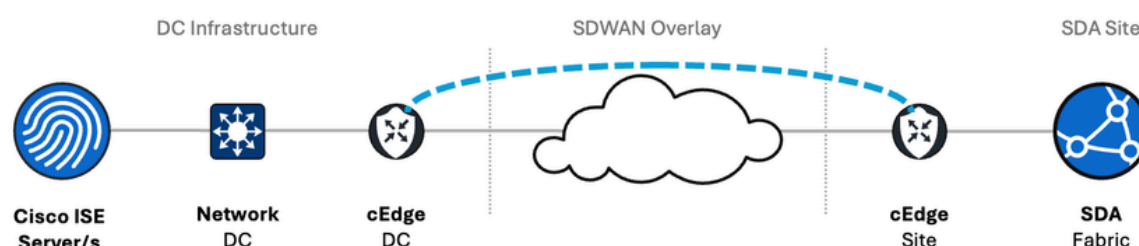
考虑常见的企业拓扑：

- 思科ISE服务器：部署在连接到数据中心网络基础设施的集中式数据中心(DC)或区域中心。
- 数据中心基础设施：包括ISE服务器所连接的DC核心或汇聚交换机。
- SD-WAN重叠：DC思科边缘路由器通过底层传输网络(例如，互联网、MPLS)与远程

SDA站点上的思科边缘路由器建立SD-WAN隧道（通常为IPsec）。

- SDA站点：远程站点思科边缘路由器连接到本地SDA交换矩阵，包括交换矩阵边缘节点、边界节点、无线局域网控制器(WLC)，并最终连接到终端。

示例拓扑



练习1:MTU缺口 — SDA边界到SD-WAN边缘

Cisco SDA设计原则（通常通过LAN Automation实施）可在所有交换矩阵设备上推广全园区9100字节（巨帧）的MTU。这包括Catalyst 9000系列边界节点，并确保以太网巨帧在交换矩阵内有效地传输。因此，SDA边界节点上的第3层或SVI切换接口默认为此更大的MTU。

相反，SD-WAN边缘设备（例如Catalyst 8000系列）通常默认为1500字节的接口MTU。这是连接外部网络（如Internet服务提供商[ISP]）的接口的标准，在这些网络中，巨型帧支持不常见或未启用。

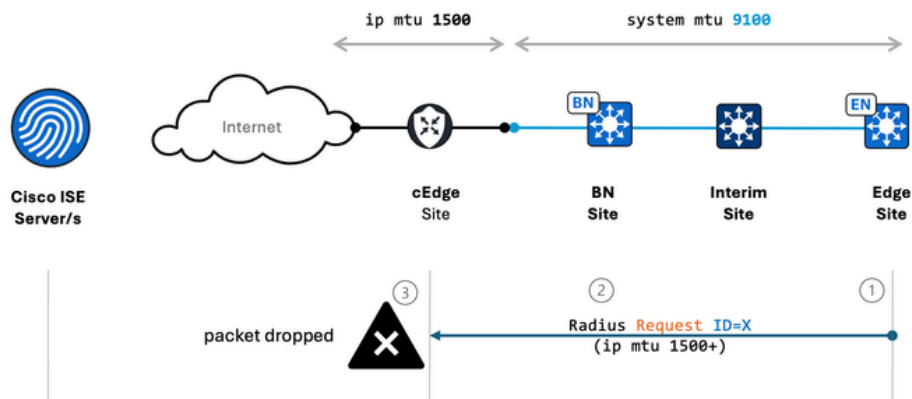
这种差异直接导致了潜在的失败：尝试将大于1500字节的IP数据包发送到其接收接口配置为1500字节MTU的SD-WAN边缘的SDA边界。

这种类型的MTU不匹配是SDA部署中的常见缺陷，在配置过程中通常很容易被忽略。更具有挑战性的是，与如何在运行Cisco IOS-XE®的Catalyst 9000交换机上生成RADIUS请求相关的某些行为会导致仅在特定和关键条件下才会出现这些问题。

例如，在会话管理器后台守护程序(SMD)进程处理的最终用户身份验证过程中生成的RADIUS请求被硬编码为以1396字节的数据包分段。相反，检索TrustSec策略（例如安全组访问控制列表[SGACL]）所涉及的RADIUS请求是由思科互联网络操作系统后台守护程序(IOSd)子组件生成的。这些是MTU感知，可避免对数据包进行分段，除非数据包的大小超过系统MTU（通常高达9100字节）。

因此，与MTU不匹配相关的问题仅在使用Cisco TrustSec(CTS)下载策略时才会变得明显。此外，在用户身份验证期间，SDA边缘设备下载的基于角色的访问控制列表(RBACL)集可能因其他标记已经存在哪些SGACL策略而异。实际上，交换机仅下载策略集的非重叠部分。

这些行为共同可能导致不可预知和不一致的结果，从无声故障到不完整策略下载，具体取决于SGACL策略的大小、当前系统条件，并最终导致路径上的MTU失调。



SDA边界通过SD-WAN边缘向ISE转发大型RADIUS数据包(例如，1600字节)，情况如下：

1. SDA边界及其9100 MTU接口发送1600字节的IP数据包。
2. SD-WAN思科边缘路由器在其1500 MTU接口上收到此数据包。
3. 但是，如果未在这些RADIUS数据包上设置Do Not Fragment(DF)位，SD-WAN思科边缘路由器通常可以在入口时丢弃它们，因为它们相对于已配置的接口MTU“过大”。它不会进入IP转发逻辑的阶段，在该阶段它可以考虑对它们进行分段（如果DF位允许）。

这种静默丢弃会导致重大故障排除问题，特别是因为问题是定向的（SDA到SD-WAN/ISE）。

类似的MTU不匹配可能发生在数据中心(DC)核心或枝叶交换机，这些交换机通常配置为支持巨型帧(例如，MTU 9000+)，以提高内部DC流量效率。但是，如果流量转移到配置了标准MTU(例如，1500字节)的SD-WAN DC思科边缘路由器路由器的面向LAN的接口，这种不匹配可能会导致分段或丢包，尤其是从DC网络流入SD-WAN交换矩阵的流量。

挑战1的解决方案：

将SDA边界切换接口（物理或SVI）上的IP MTU与对等的SD-WAN思科边缘路由器接口（通常为1500字节）对齐。

配置示例（在SDA边界节点上）：

```
<#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
!
```

重要注意事项：Catalyst 9000边界上的分段

Catalyst 9000系列交换机作为SDA边界节点，支持硬件数据平面中的本地IP数据包的IP分段。将切

换接口上的ip mtu减少到1500不会导致性能降低，因为源于或经过所需边界的数据流的基于软件的分片会降低性能。交换机在离开此特定接口之前对大于1500字节的IP数据包进行有效的分段（如果DF位已清除），而不会传送到CPU。

但是，请注意，Catalyst 9000交换机通常不支持VXLAN封装流量的分段。此限制对重叠流量至关重要，但不会影响所描述的RADIUS身份验证场景，因为SDA边界和外部ISE之间的RADIUS通信通常发生在底层（本地IP路由）内。（VXLAN重叠的MTU注意事项是一个单独且复杂的主题，详见相关的思科SDA设计指南）。

SDA边界的主动式MTU与SD-WAN的协调思科边缘路由器切换至关重要。

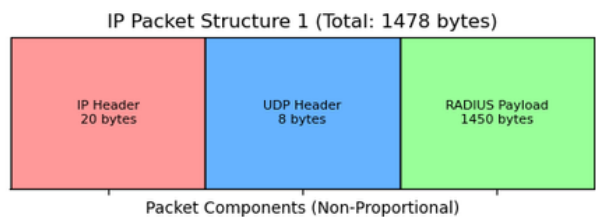
练习2:TINC Garbage Ltd.通过SD-WAN重叠的MTU Squeeze - ISE流量

即使单个物理接口(例如ISE网络接口卡(NIC)、交换机端口或路由器接口)设置为标准的1500字节IP MTU，SD-WAN重叠本身也会带来封装开销。此开销会消耗1500字节限制的一部分，从而降低可用于原始IP数据包的有效MTU(从ISE的角度看是“负载”)。

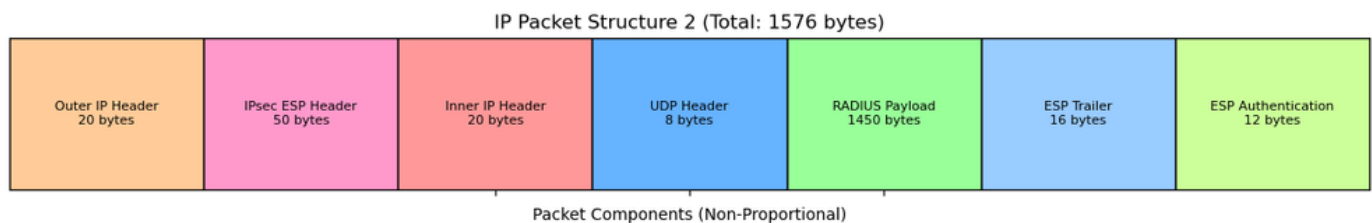
数据包结构和封装开销：

当从ISE服务器(例如，RADIUS Access-Accept数据包)发送至SDA站点中的网络接入设备(NAD)时，数据包会穿越SD-WAN重叠并封装。常见的封装堆栈涉及隧道模式下的IPsec，可能通过UDP进行NAT穿越(NAT-T)。

- 来自ISE的原始数据包（内部数据包）：
例如，具有1450字节负载+ 8B UDP + 20B内部IP的RADIUS数据包= 1478字节。

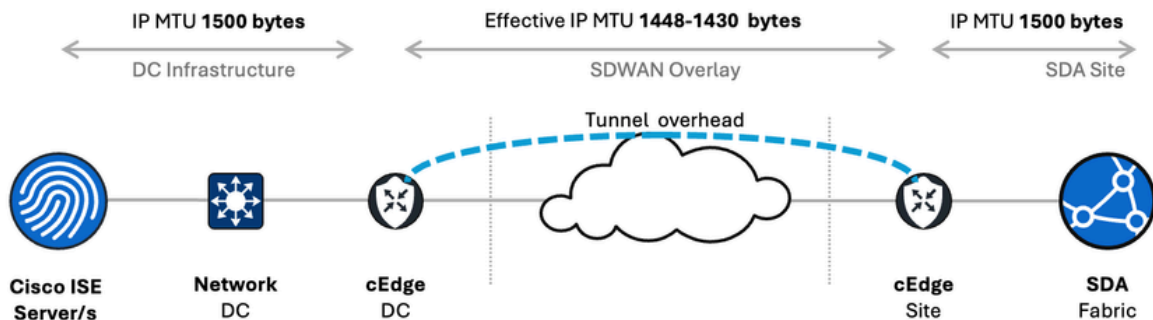


- 考虑隧道模式下的IPsec ESP，可能使用用于NAT-T的UDP封装：



- 总开销可能因特定IPsec密码、身份验证机制和其他重叠功能(如GRE（如果使用）)而异。典型计算：
 - 外部IP报头(IPv4):20 bytes

- UDP报头 (如果用于NAT-T的ESP over UDP) : 8 个字节
- ESP 报头:~8字节
- ESP IV(例如, 适用于AES-CBC):~16字节 (如果适用)
- ESP身份验证(例如, HMAC-SHA256被截断):~12-16字节
- 常见IPsec估计开销: ~52-70字节 (可能更高, 使用所有选项最多可达~80字节或更多)。



如果物理链路MTU为1500字节, 来自ISE的原始IP数据包的可用负载MTU将变为: 1500字节 — SD-WAN开销。

例如, $1500 - 70 = 1430$ 字节。

数据包超过有效MTU时的行为:

1. ISE生成数据包 (DF位异常):

- 默认情况下, ISE设备的底层Linux操作系统在IP报头中为小于或等于其配置的接口IP MTU(例如, 1500字节)的所有数据包设置Do Not Fragment(DF)。
- 此DF位的用途: ISE (通过其OS) 主动设置DF位, 主要是利用路径MTU发现 (PMTUD)流程, 稍后将对此进行介绍。这样, 如果ISE的PMTU小于其自己的接口MTU, ISE便可以动态获知到目标的实际PMTU。
- 大于接口MTU的数据包行为: 如果ISE需要发送的IP数据包大于其配置的接口IP MTU, 则行为取决于其Linux操作系统。通常, OS可以在传输之前对数据包进行分段, 并清除这些生成的分段上的DF位 (设置DF=0)。此分段是操作系统级功能, 不直接由ISE应用代码本身驱动。
- 与网络设备的关键区别: ISE的此默认行为 (设置DF=1, 即使对于适合其接口MTU的非分段数据包) 与许多传统网络设备 (路由器、交换机) 有很大不同。网络设备通常不对它们发起或转发的数据包设置DF位, 除非明确配置为这样做, 或者如果被转发的数据包已设置DF位, 或者为特定协议设置DF位。如果数据包超过下一跳MTU (且DF=0), 它们通常默认允许分段。
- 故障排除复杂性: 这种不对称性(ISE到NAD流量默认情况下通常为DF=1, 而NAD到ISE流量可为DF=0 (除非需要因某种原因进行设置)) — 会在故障排除期间引入额外的复杂性。工程师可以根据流量方向观察不同的分段行为和PMTUD交互。

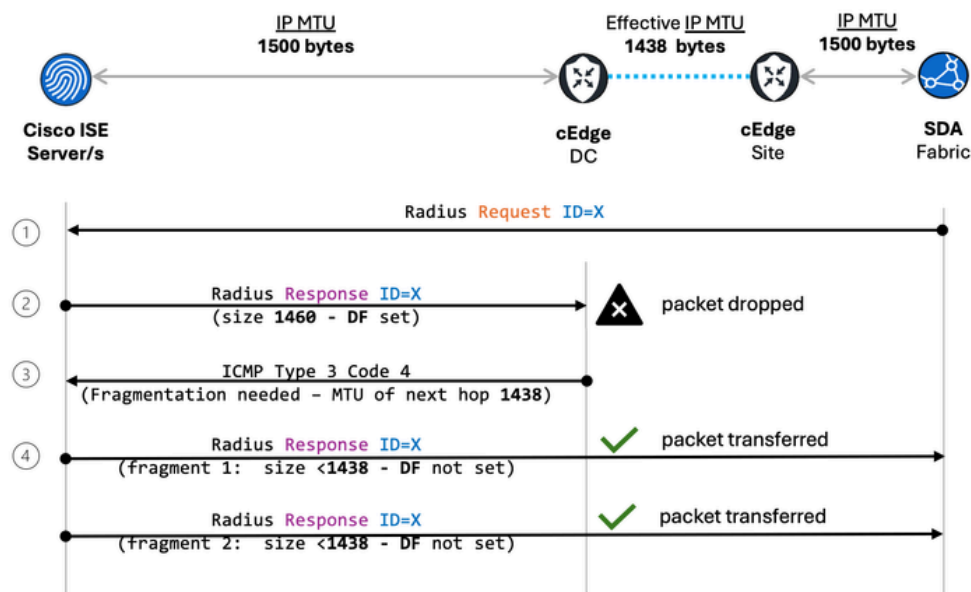
2. 数据包到达入口Cisco Edge路由器(DC):DC Cisco Edge路由器从ISE接收IP数据包。

3. 思科边缘路由器的封装和MTU检查: 思科边缘路由器尝试为SD-WAN隧道封装数据包。

- 如果原始数据包大小为SD-WAN封装开销超过思科边缘路由器的出站物理接口MTU(例如，1500字节)，并且在来自ISE的原始（内部）数据包上设置DF位，则思科边缘路由器不得对内部数据包进行分段。
- Cisco Edge路由器应该丢弃数据包。
- 关键是，思科边缘路由器还应将ICMP“Destination Unreachable - Fragmentation Needed and DF bit set”（类型3，代码4）消息发送回源(ISE)，指示下一跳的MTU（隧道的有效MTU）。

4. 路径MTU发现(PMTUD)进程：收到此ICMP“需要分段”消息后，ISE（源操作系统）应减少该特定目标路径的PMTU估计值。它会缓存此信息，并以适合新发现的PMTU的小数据包重新发送数据。

PMTUD流程图：



PMTUD通信中断的位置：

PMTUD在理论上是稳健的，但在实践中可能失败：

- **ICMP过滤**：中间防火墙或安全策略通常会阻止ICMP消息，阻止“需要分段”消息到达ISE。
- **思科边缘路由器上的控制平面策略(CoPP)**：Cisco边缘路由器使用CoPP保护其CPU。生成ICMP错误消息是一项控制平面任务。在负载较重或存在许多超大数据包的情况下，CoPP可以速率限制或丢弃ICMP生成。ISE永远不会收到反馈。
- **静默丢弃**：如果ISE未收到ICMP“需要分段”消息，它仍然不知道路径限制。它继续发送设置了DF位的大型数据包，导致入口思科边缘路由器以静默方式丢弃这些数据包。这会导致应用层超时和重传(例如RADIUS)。
- **对ISE服务的影响**：大型RADIUS Access-Accept数据包（携带dACL、大量AVP和SGT信息）特别容易受影响。表现包括：
 - 间歇或完全身份验证失败。
 - 未收到正确网络访问策略或SGT的终端。

- ISE和NAD之间的策略同步不完整或失败。

挑战2的解决方案：主动ISE IP MTU配置

由于PMTUD不可靠，因此主动方法最适用于ISE等关键服务。在ISE的网络接口上配置IP MTU，使其值可以安全容纳最大预期的SD-WAN重叠开销。这可以确保ISE不会发起IP数据包（设置了DF位），这些数据包本身太大而无法通过SD-WAN重叠，而无需中间设备分段（如果DF=1，则禁止分段）。

计算和设置建议的ISE IP MTU:

1. 建立基本物理MTU：对于路径上的标准以太网接口，通常为1500字节。
2. 确定最大SD-WAN封装开销：
 - 准确计算或保守地估计您的特定SD-WAN重叠（IPsec、GRE、VXLAN、MPLSoGRE等）引入的总开销。有关您选择的协议和选项的精确数字，请参阅供应商文档。

组件	开销示例（字节）	备注
基本物理MTU	1500	物理链路上的标准以太网
减：SD-WAN开销		
外部IP报头(IPv4)	20	
UDP报头（用于NAT-T）	8	如果ESP封装在UDP中
ESP报头	~8-12	
ESP IV(例如,AES-CBC)	~16	因加密算法而异
ESP身份验证(例如，SHA256)	~12-16	因身份验证算法而异（例如，有些为96位）
其他重叠（GRE等）	变量	添加SD-WAN封装堆栈的一部分
总估计开销	~68 - 80个以上字节	用于部署的所有相关组件的总和
有效路径MTU	~1432 - 1420字节	基本物理MTU — 估计总开销

3. 建议的ISE IP MTU配置：
 - 采用计算出的有效路径MTU（例如，示例中的1420字节）。
 - 减去额外的安全余量（例如，20-70字节），以说明次要的、未入帐的L2报头或提供缓冲区。
 - Cisco SD-WAN等解决方案可以为每个站点到站点隧道单独执行路径MTU(PMTU)发现。此机制每20分钟自动运行一次，以根据每个站点的当前传输条件测试和动态调整隧道的IP MTU。因此，不同站点的MTU值可能不同，并且会随时间而变化。
 - 在此类场景中，ISE接口的一般安全和建议的IP MTU介于1350和1400字节之间

1350字节的IP MTU通常是一个非常稳健的起点

ISE配置（通过CLI的示例）：

此命令在思科ISE设备CLI上为每个相关网络接口执行。

```
<#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项：

- 需要服务重启：对ISE接口应用ip mtu命令后，将提示用户重新启动ISE应用服务。这是一个影响服务的变更，必须在计划维护时段内进行安排。有关程序详细信息，请参阅官方Cisco ISE文档。
- 应用于所有ISE节点：此IP MTU调整必须一致地应用于部署中与SD-WAN上的NAD通信的所有ISE节点(主PAN、辅助PAN、策略服务节点(PSN))。MTU设置不一致会导致无法预测的行为。
- 全面测试：在生产中实施之前，请在实验室或试运行部署中严格测试此更改。使用各种工具(如ping)验证端到端MTU处理，这些工具具有不同的数据包大小和DF位设置：
 - 基于Linux的系统：

```
ping
```

```
-s
```

```
-M do
```

(注：-s指定ICMP负载大小。总IP数据包大小=负载+ 8B ICMP Hdr + 20B IP Hdr (对于IPv4))

- Windows 窗口版本:

```
ping
```

```
-f -l
```

(注：-l指定ICMP负载大小。)

- Cisco IOS/Cisco IOS-XE®

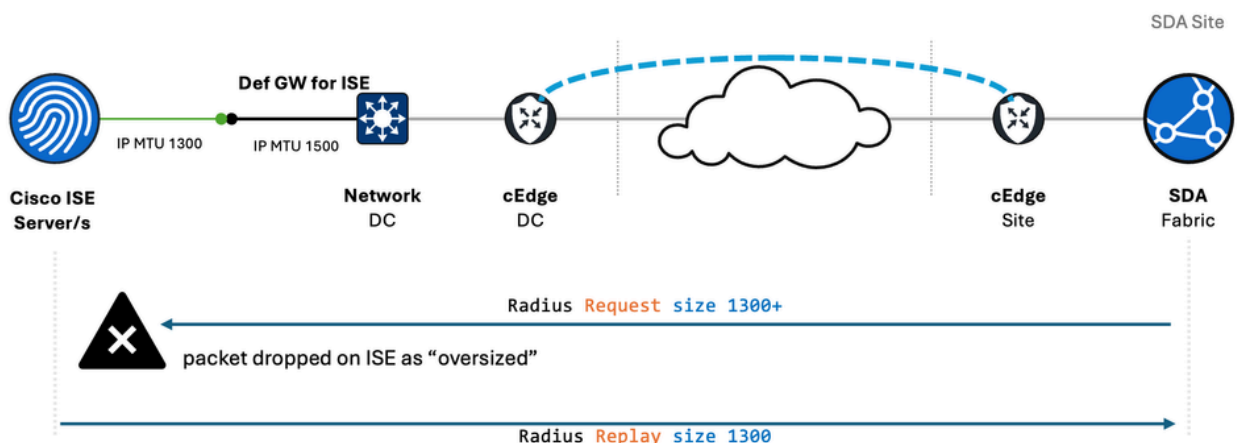
ping

size

df-bit

- ISE第一个路由点 — 调整ISE接口上的IP MTU值时，请确保数据中心中的第一个路由点（尤其是与ISE子网关联的第3层接口）也配置了相同的IP MTU值。这有助于防止类似于练习1中所述的情况，其中MTU不匹配导致ISE将传入数据包视为过大并丢弃它们。
例如，如果ISE接口的MTU减少(例如，1300)，但第一个路由点仍使用默认MTU 1500进行配置，则发送到ISE的大于1300字节但小于1500字节的数据包不会分段，并由ISE丢弃，如挑战1中所示。
此外，确保第一个路由点能够在需要时执行分段，而且这样做不会导致性能降低。
- 更新整个传输路径及两个方向的MTU — 更新ISE上的IP MTU设置时，必须考虑整个传输路径及两个方向的MTU。如果ISE上配置的MTU值与第一跳网关的第3层接口上的MTU不一致，可能会出现类似的问题，如练习#1中所述。

例如，如果ISE MTU减少到1300字节，而默认1500字节MTU仍配置在默认网关上，则网络设备通常生成的1300到1500字节大小的数据包可能会被ISE丢弃为过大。



为避免此问题，请始终确保ISE上的MTU更改在第一跳网关上镜像，并且最好反映在同一第3层网段内的所有终端主机上。这有助于保持端到端MTU一致性，并防止意外丢包。

结论

按照SD-WAN封装和SDA边界的MTU调整所施加的有效传输层MTU限制来调整Cisco ISE服务器上的IP MTU设置与SD-WAN的MTU调整思科边缘路由器切换不仅是一项建议，而且是在现代分段企业网络中确保AAA服务的稳定性、可靠性和性能的关键前提条件。虽然路径MTU发现是一种重要的机制，但其实际有效性可能会受到SD-WAN环境中ICMP过滤或控制平面策略等因素的影响。

通过主动在ISE上配置更少的IP MTU(例如，1350-1400字节)，网络架构师和工程师可以显著降低与MTU相关的丢包风险，从而获得更可预测且恢复力更强的网络运营。这在思科SDA部署中尤其重要，ISE在思科SDA部署中协调复杂的微分段和动态策略实施，这些实施通常依赖于成功交付潜在的大型控制平面消息。在所有ISE节点之间进行勤奋规划、全面测试和一致配置是成功且无故障部署的关键。

标准和参考

要更深入地了解，请参阅官方标准和思科文档：

RFC:

- RFC 1191：路径MTU发现
- RFC 791:Internet协议(IP) — 定义IP报头，包括不分段(DF)位。
- RFC 8200：IPv6规范（如果使用IPv6则相关，包括类似的PMTUD概念）。
- RFC 4459：网内隧道(VPN)的MTU和分段问题 — 直接解决VPN环境中的常见MTU问题。

思科文档：

- Cisco SDA设计和部署指南：有关交换矩阵MTU建议和边界节点配置的信息。
- Cisco SD-WAN设计和配置指南：有关封装开销、隧道接口MTU以及SD-WAN交换矩阵中的PMTUD注意事项的详细信息。
- Cisco Catalyst 9000系列交换机配置指南：有关MTU设置和分段功能的平台特定详细信息。
- 思科身份服务引擎(ISE)管理员和CLI指南：有关接口配置的信息，包括ip mtu命令和服务重新启动影响。

关于此翻译

思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。