

Découverte MTU de chemins IP et DLSw

Contenu

[Introduction](#)

[Avant de commencer](#)

[Conventions](#)

[Conditions préalables](#)

[Components Used](#)

[Informations générales](#)

[DLSw avec PMTD](#)

[Vérification de PMTD pour DLSW](#)

[Informations connexes](#)

[Introduction](#)

La suite de protocoles IBM, DLSw, STUN et BSTUN établissent un canal de session IP d'un routeur à l'autre. Le protocole TCP est généralement utilisé comme méthode de transport entre les routeurs en raison de sa fiabilité. Ce document fournit des informations sur la capacité du protocole TCP à détecter dynamiquement la plus grande unité de transmission maximale pouvant être utilisée sur le canal de session, ce qui réduit la fragmentation et optimise l'efficacité.

[Avant de commencer](#)

[Conventions](#)

Pour plus d'informations sur les conventions des documents, référez-vous aux [Conventions utilisées pour les conseils techniques de Cisco](#).

[Conditions préalables](#)

Aucune condition préalable spécifique n'est requise pour ce document.

[Components Used](#)

Ce document n'est pas limité à des versions de matériel et de logiciel spécifiques.

Les informations présentées dans ce document ont été créées à partir de périphériques dans un environnement de laboratoire spécifique. All of the devices used in this document started with a cleared (default) configuration. Si vous travaillez dans un réseau opérationnel, assurez-vous de bien comprendre l'impact potentiel de toute commande avant de l'utiliser.

Informations générales

Path MTU Discovery (PMTD), décrit dans la RFC 1191, spécifie que la taille d'octet par défaut d'un paquet IP est de 576. Les parties IP et TCP de la trame occupent 40 octets, laissant 536 octets comme charge utile de données. Cet espace est appelé taille de segment maximale ou MSS. La section 3.1 de la RFC1191 spécifie qu'un MSS plus grand peut être négocié, et c'est exactement ce que l'exécution de la commande **ip tcp path-mtu-discovery** fait dans un routeur Cisco. Lorsque cette commande est configurée et qu'une session TCP est démarrée, le paquet SYN du routeur contient une option TCP spécifiant un MSS plus grand. Ce MSS plus grand est le MTU de l'interface sortante moins 40 octets. Si le MTU de l'interface de sortie est de 1 500 octets, le MSS annoncé est de 1 460. Si l'interface de sortie a une MTU plus grande, par exemple, Frame Relay avec une MTU de 4 096 octets, le MSS sera de 4 096 octets moins 40 octets d'informations IP et sera affiché dans la sortie de la commande **show tcp** (le segment de données max est 4 056 octets).

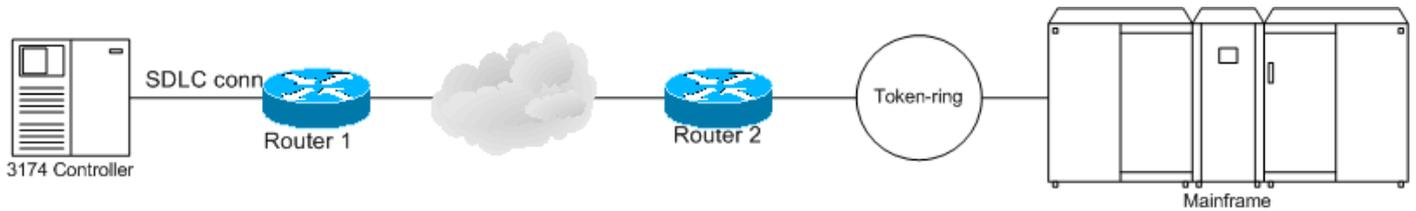
La configuration de PMTD sur le routeur n'a aucun effet sur les sessions TCP existantes déjà établies depuis/vers le routeur. PMTD a été introduit dans le niveau IOS 11.3.5T et dans les versions ultérieures d'IOS, il est devenu une commande facultative. Avant IOS 11.3(5)T, il était activé par défaut. En outre, PMTD est automatique lorsque les adresses IP se trouvent dans le même sous-réseau, indiquant qu'elles sont directement connectées sur le même support.

Les deux routeurs doivent être configurés pour que PMTD fonctionne correctement. Lorsque les deux routeurs sont configurés, le SYN d'un routeur à l'autre contient la valeur TCP facultative annonçant le MSS le plus élevé. Le SYN retourné annonce ensuite la valeur MSS la plus élevée. Ainsi, les deux parties font de la publicité à l'autre ils peuvent accepter un MSS plus grand. Si la commande **ip tcp path-mtu-discovery** est configurée sur un seul routeur, le routeur 1, annonce le MSS plus grand et le routeur 2 peut donc envoyer une trame de 1 460 octets au routeur 1. Le routeur 2 n'annoncera jamais le MSS le plus grand, et le routeur 1 est donc verrouillé pour envoyer les valeurs par défaut.

DLSw avec PMTD

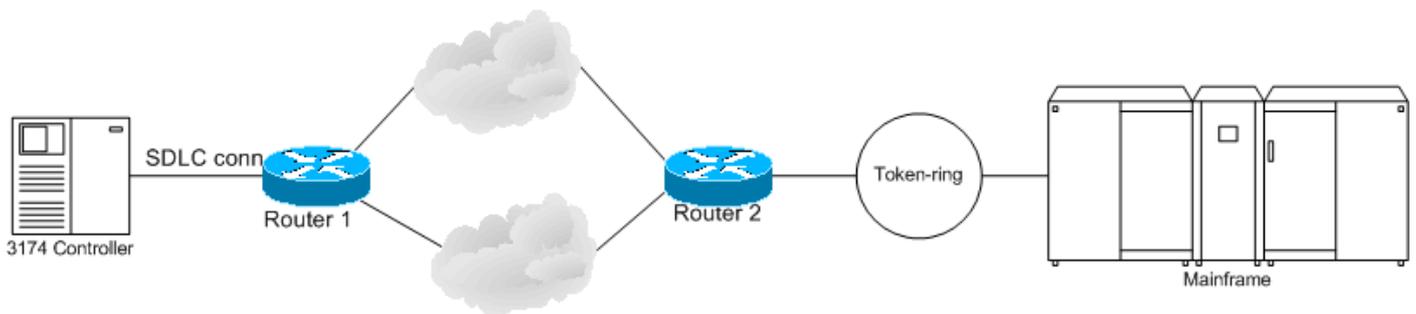
Dans la suite IBM, DLSw, STUN et BSTUN peuvent être chargés de transporter de grandes quantités de données sur une session TCP d'un routeur à l'autre. Il peut être important et extrêmement utile d'implémenter PMTD, surtout si l'on considère qu'elle a été activée par défaut dans les niveaux 11.2 et IOS précédents. Selon le RFC, la trame la plus importante est de 576 octets par défaut, moins 40 octets pour l'encapsulation IP/TCP. DLSw utilise 16 octets supplémentaires pour l'encapsulation. Les données réelles qui sont transportées, à l'aide du MSS par défaut, sont de 520 octets. DLSw peut également transporter deux paquets LLC2 (Logical Link Control 2) différents dans une trame TCP. Si deux stations de travail envoient chacune une trame LLC2, DLSw peut transporter les deux trames LLC2 vers le pair distant DLSw en une seule trame. En disposant d'un MSS plus grand, les pilotes TCP peuvent prendre en charge ce schéma de piggyback. Voici trois scénarios principaux pour illustrer la valeur de la commande **path-mtu-discovery**.

Scénario 1 - Frais généraux indésirables



Les périphériques SDLC sont généralement configurés pour un maximum de 265 ou 521 octets de données dans chaque trame. Si la valeur est 521 et que le 3174 envoie au routeur 1 une trame SDLC de 521 octets, le routeur 1 crée deux trames TCP pour les transmettre au routeur homologue DLSW Router 2. La première trame contient 520 octets de données, 16 octets d'informations DLSw et 40 octets d'informations IP pour un total de 576 octets. Le deuxième paquet contient 1 octet de données, 16 octets d'informations DLSw et 40 octets d'informations IP. Lorsque le PMTD est utilisé et que le routeur 2 supposait un MTU de 1 500 octets pour obtenir un MSS de 1 460, le routeur 1 a été informé que le routeur 2 pouvait recevoir 1 460 octets de données. Le routeur 1 envoie les 521 octets de données SDLC au routeur 2 en un seul paquet avec 16 octets d'informations DLSw et 40 octets d'informations IP. Puisque DLSw est un événement commuté de processus, en utilisant PMTD, l'utilisation du CPU pour traiter cette trame SDLC unique a été réduite de moitié. En outre, le routeur 2 n'a pas à attendre que le deuxième paquet assemble la trame LLC2. Avec, PMTD activé, le routeur 2 reçoit l'intégralité du paquet et peut ensuite supprimer les informations IP et DLSW du paquet et l'envoyer au routeur 3745 sans délai.

Scénario 2 - Délai des paquets hors commande

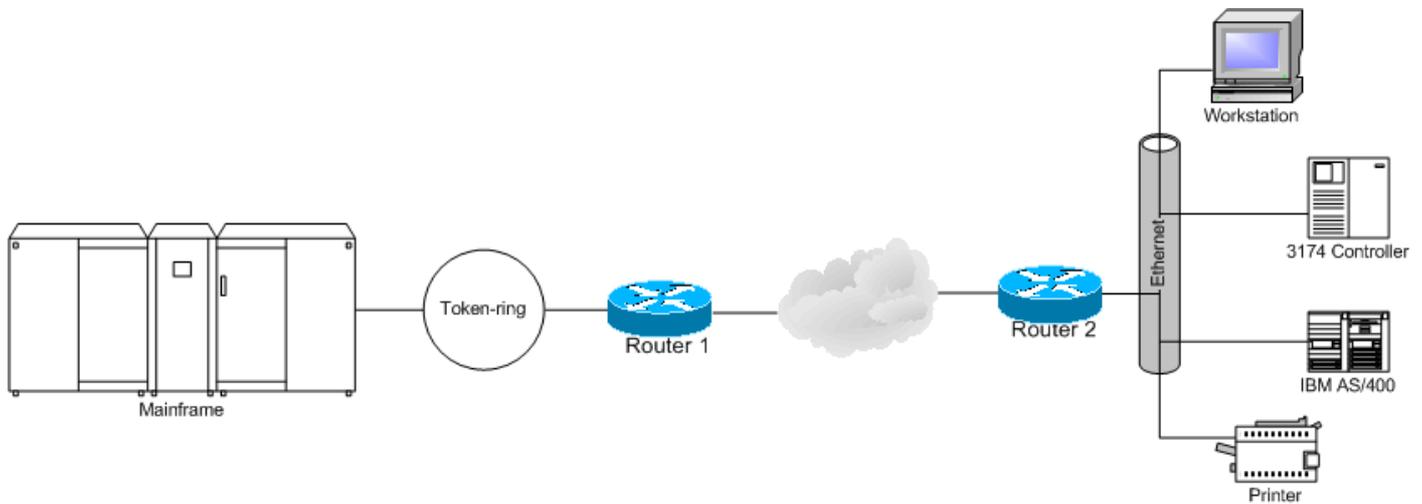


Dans ce scénario, deux nuages IP sont disponibles avec des métriques égales pour l'équilibrage de charge ou la redondance. L'inactivation de PMTD peut ralentir considérablement le DLSw. Sans PMTD activé, le routeur 1 doit assembler la trame de 521 octets en deux paquets TCP : un avec 520 octets de données et un autre avec 1 octet de données. Si le premier paquet traverse le cloud IP supérieur, il est fort probable qu'il arrive après le second paquet si le second paquet est envoyé via le cloud IP à coût égal inférieur. Cela génère ce que l'on appelle un paquet hors service. La capacité à gérer ce problème est inhérente à la capacité du protocole IP/TCP. Les paquets en panne sont stockés dans la mémoire en attendant l'arrivée de l'intégralité du flux, puis leur réassemblage. Les paquets en panne ne sont pas rares. Cependant, toutes les tentatives de réduction doivent être faites car cet événement utilise la mémoire et les ressources du processeur. Une grande quantité de commandes sortantes peut provoquer un retard important au niveau TCP. Si la session de couche3/DLSw est retardée, la session LLC2/SDLC qui est transportée sur ce DLSw en pâtira par la suite. Si PMTD est activé dans ce scénario, une trame unique de 521 octets est envoyée dans une trame TCP sur l'un ou l'autre des nuages IP. Le routeur récepteur n'a besoin que d'une mémoire tampon et d'une désencapsulation d'une trame TCP.

PMTD n'a aucune relation avec la plus grande trame annoncée de station d'extrémité à station d'extrémité dans les environnements SNA. Cela inclut la trame la plus grande (LF) du champ RIF (Routing Information Field) sur Token Ring. Le protocole PMTD détermine strictement la quantité de données pouvant être encapsulées dans une trame TCP. LLC2 et SDLC n'ont pas les paquets

de fragment de capacité, mais IP/TCP le fait. Une trame SNA de grande taille peut être segmentée lorsqu'elle est encapsulée dans TCP. Ces données sont décapsulées au niveau du routeur DLSw distant, puis présentées de nouveau comme des données SNA non fragmentées.

Scénario 3 : connectivité et débit LLC2 plus rapides



Dans ce scénario, le 3174 et la station de travail ont des sessions via le 3745 TIC vers le mainframe, si les deux périphériques envoient des données destinées à l'hôte, il est possible que TCP puisse placer les deux trames LLC2 dans un paquet. Sans PMTD, cela n'est pas possible si les données combinées des deux trames sont de 521 octets ou plus. Dans ce cas, le logiciel TCP devra envoyer chaque paquet séparément. Par exemple, si le 3174 et la station de travail envoient une trame à peu près au même moment et que chacun de ces paquets a 400 octets de données, le routeur reçoit et met en mémoire tampon chaque trame. Le routeur doit maintenant encapsuler chacun de ces flux de données de 400 octets dans des paquets TCP séparés pour transmission à l'homologue.

Avec PMTD activé et en supposant un MSS de 1460, le routeur reçoit et met en mémoire tampon les deux paquets LLC2. Il peut maintenant encapsuler les deux dans un seul paquet. Ce paquet TCP contiendra 40 octets d'informations IP, 16 octets d'informations DLSw pour le premier couplage de circuits DLSw, les 400 octets de données, 16 octets de données supplémentaires pour le deuxième couplage de circuits DLSw et les 400 autres octets de données. Ce scénario particulier utilise deux périphériques et donc deux circuits DLSw. Le protocole PMTD permet à DLSw d'évoluer vers un nombre plus élevé de circuits DLSw plus efficacement. De nombreux réseaux en étoile nécessitent des centaines de sites distants, chacun équipé d'un ou de deux périphériques SNA, qui s'appairent à un routeur de site central connecté à un OSA ou à un FEP fournissant l'accès aux applications hôtes. Le protocole PMTD permet aux protocoles TCP et DLSw de s'adapter à des besoins plus importants sans surutiliser le processeur et les ressources de mémoire du routeur, ainsi qu'un transport plus rapide.

Note : Un bogue logiciel a été détecté à la fin de la version 12.1(5)T et résolu dans la version 12.2(5)T où PMTD ne fonctionnait pas sur un tunnel VPN. Ce défaut logiciel est [CSCdt49552](#) (clients [enregistrés](#) uniquement).

Vérification de PMTD pour DLSW

Émettez la commande **show tcp**.

```
havoc#show tcp
```

```
Stand-alone TCP connection to host 10.1.1.1
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 30.1.1.1, Local port: 11044
Foreign host: 10.1.1.1, Foreign port: 2065

Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0 bytes)

TCP driver queue size 0, flow controlled FALSE
```

```
Event Timers (current time is 0xA18A78):
```

Timer	Starts	Wakeups	Next
Retrans	3	0	0x0
TimeWait	0	0	0x0
AckHold	0	0	0x0
SendWnd	0	0	0x0
KeepAlive	0	0	0x0
GiveUp	2	0	0x0
PmtuAger	0	0	0x0
DeadWait	0	0	0x0

```
iss: 3215333571  snduna: 3215334045  sndnxt: 3215334045      sndwnd: 20007
irs: 3541505479  rcvnxt: 3541505480  rcvwnd:      20480  delrcvwnd: 0
```

```
SRTT: 99 ms, RTTO: 1539 ms, RTV: 1440 ms, KRTT: 0 ms
minRTT: 24 ms, maxRTT: 300 ms, ACK hold: 200 ms
Flags: higher precedence, retransmission timeout
```

```
Datagrams (max data segment is 536 bytes):
```

```
Rcvd: 30 (out of order: 0), with data: 0, total data bytes: 0
Sent: 4 (retransmit: 0, fastretransmit: 0), with data: 2, total data bytes: 473
```

Cet affichage est identifié comme une session TCP DLSw car l'un des ports de la session TCP est 2065. Près du bas de la sortie est le segment de données max 536 octets. Cette valeur indique que le routeur homologue DLSw distant 10.1.1.1 n'a pas la commande **ip tcp path-mtu-discovery** configurée. La valeur de 536 octets représente déjà les 40 octets d'informations IP dans une trame IP. Cette valeur de 536 octets ne prend pas en compte les 16 octets d'informations DLSw qui seraient ajoutés à un paquet TCP transportant du trafic SNA.

Avec la commande **ip tcp path-mtu-discovery** configurée, le segment de données max est maintenant 1460. En outre, la sortie de la commande **show tcp** indique que le **chemin mtu est compatible** immédiatement avant l'instruction de **segment de données max**. L'interface de sortie a un MTU de 1 500 octets. MTU équivaut à 1 500 octets moins 40 octets d'informations IP sont de 1 460 octets. DLSw utilisera 16 octets supplémentaires. Par conséquent, une trame LLC2 ou SDLC de 1 444 octets maximum peut être transmise dans une trame TCP.

```
havoc#show tcp
```

```
Stand-alone TCP connection to host 10.1.1.1
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 30.1.1.1, Local port: 11045
Foreign host: 10.1.1.1, Foreign port: 2065

Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0 bytes)

TCP driver queue size 0, flow controlled FALSE
```

```
Event Timers (current time is 0xA6DA58):
```

Timer	Starts	Wakeups	Next
Retrans	4	0	0x0
TimeWait	0	0	0x0
AckHold	1	0	0x0
SendWnd	0	0	0x0
KeepAlive	0	0	0x0
GiveUp	3	0	0x0
PmtuAger	0	0	0x0
DeadWait	0	0	0x0

iss: 3423657490 snduna: 3423657976 sndnxt: 3423657976 sndwnd: 19995
irs: 649085675 rcvnxt: 649085688 rcvwnd: 20468 delrcvwnd: 12

SRTT: 124 ms, RTTO: 1405 ms, RTV: 1281 ms, KRTT: 0 ms
minRTT: 24 ms, maxRTT: 300 ms, ACK hold: 200 ms
Flags: higher precedence, retransmission timeout, path mtu capable

Datagrams (max data segment is 1460 bytes):
Rcvd: 5 (out of order: 0), with data: 1, total data bytes: 12
Sent: 6 (retransmit: 0, fastretransmit: 0), with data: 3, total data bytes: 485

[Informations connexes](#)

- [Notes techniques Compatible Systems : Fragmentation IP et découverte de chemin MTU avec VPN](#)
- [Support technique - Cisco Systems](#)