

# Rilevamento MTU percorso IP e DLSw

## Sommario

[Introduzione](#)

[Operazioni preliminari](#)

[Convenzioni](#)

[Prerequisiti](#)

[Componenti usati](#)

[Premesse](#)

[DLSw con PMTD](#)

[Verifica del PMTD per DLSW](#)

[Informazioni correlate](#)

## [Introduzione](#)

La suite di protocolli IBM DLSw, STUN e BSTUN stabiliscono un pipe di sessione IP da un router all'altro. Il protocollo TCP viene in genere utilizzato come metodo di trasporto tra router a causa della sua affidabilità. Questo documento offre informazioni sulla capacità del protocollo TCP di individuare dinamicamente la MTU più grande che può essere utilizzata sulla pipe di sessione, riducendo al minimo la frammentazione e massimizzando l'efficienza.

## [Operazioni preliminari](#)

### [Convenzioni](#)

Per ulteriori informazioni sulle convenzioni usate, consultare il documento [Cisco sulle convenzioni nei suggerimenti tecnici](#).

### [Prerequisiti](#)

Non sono previsti prerequisiti specifici per questo documento.

### [Componenti usati](#)

Il documento può essere consultato per tutte le versioni software o hardware.

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

## Premesse

Il rilevamento della MTU del percorso (PMTD), descritto nella RFC 1191, specifica che le dimensioni in byte predefinite di un pacchetto IP sono 576. Le parti IP e TCP del frame occupano 40 byte, lasciando 536 byte come payload dei dati. Questo spazio è noto come dimensione massima del segmento o MSS. La sezione 3.1 della RFC 1191 specifica che è possibile negoziare un valore MSS più grande. Questo è esattamente ciò che fa il comando **ip tcp path-mtu-discovery** su un router Cisco. Quando si configura questo comando e si avvia una sessione TCP, il pacchetto SYN inviato dal router contiene un'opzione TCP che specifica un valore MSS più grande. Il valore MSS più grande corrisponde alla MTU dell'interfaccia in uscita meno 40 byte. Se l'MTU dell'interfaccia in uscita è 1500 byte, il valore MSS annunciato è 1460. Se l'interfaccia in uscita ha una MTU più grande, ad esempio Frame Relay con MTU di 4096 byte, il valore MSS sarà di 4096 byte meno 40 byte di informazioni IP e verrà visualizzato nell'output del comando **show tcp** (il segmento dati massimo è di 4056 byte).

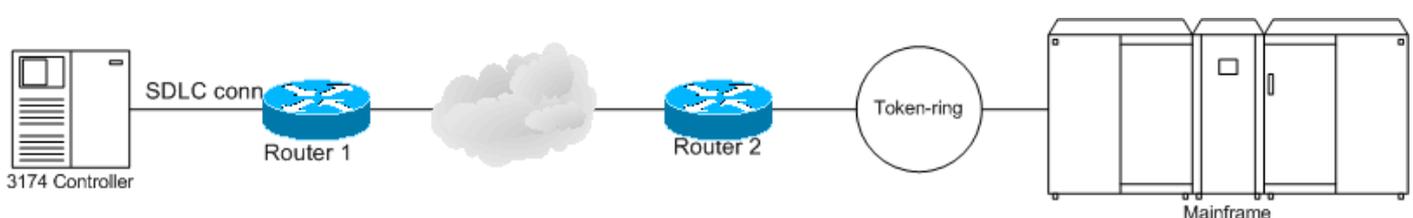
La configurazione del PMTD sul router non ha alcun effetto sulle sessioni TCP esistenti già stabilite da/verso il router. Il PMTD è stato introdotto nel livello IOS 11.3.5T e nelle versioni successive di IOS, è diventato un comando opzionale. Prima della versione IOS 11.3(5)T, questa opzione era attiva per impostazione predefinita. Inoltre, il PMTD è automatico quando gli indirizzi IP si trovano nella stessa subnet, a indicare che sono collegati direttamente allo stesso supporto.

Per il corretto funzionamento del PMTD, è necessario configurare entrambi i router. Quando sono configurati entrambi i router, il valore SYN tra un router e l'altro contiene il valore TCP facoltativo che indica il valore MSS più alto. Il valore SYN restituito annuncia quindi il valore MSS più alto. Pertanto, entrambi i dispositivi pubblicizzano l'altro e possono accettare un valore MSS più grande. Se solo un router, il router 1, ha il comando **ip tcp path-mtu-discovery** configurato, annuncerà il valore MSS più grande e quindi il router 2 potrà inviare al router 1 un frame di 1460 byte. Il router 2 non annuncerà mai il valore MSS più grande e quindi il router 1 è bloccato all'invio dei valori predefiniti.

## DLSw con PMTD

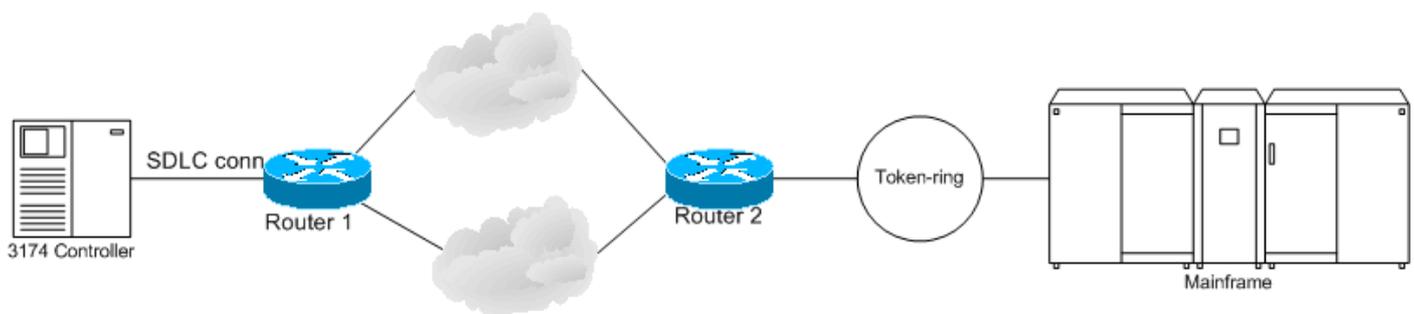
Nella suite IBM, DLSw, STUN e BSTUN possono trasportare grandi quantità di dati su una sessione TCP da un router all'altro. L'implementazione della PMTD può essere importante ed estremamente vantaggiosa, soprattutto se si considera che è stata abilitata per impostazione predefinita nei livelli 11.2 e IOS precedenti. In base all'RFC, per impostazione predefinita il frame più grande è 576 byte, meno 40 byte per l'incapsulamento IP/TCP. DLSw utilizza altri 16 byte per l'incapsulamento. Il valore MSS predefinito utilizzato per il trasporto dei dati è 520 byte. DLSw è anche in grado di trasportare due diversi pacchetti LLC2 (Logical Link Control 2) in un frame TCP. Se due workstation inviano ciascuna un frame LLC2, DLSw può trasportare entrambi i frame LLC2 al peer remoto DLSw in un frame. Se il valore MSS è maggiore, i driver TCP possono supportare questo schema di piggyback. Di seguito vengono riportati tre scenari principali per illustrare il valore del comando **path-mtu-discovery**.

### Scenario 1 - Sovraccarico indesiderato



I dispositivi SDLC sono generalmente configurati per un massimo di 265 o 521 byte di dati in ciascun frame. Se il valore è 521 e il router 3174 invia al router 1 un frame SDLC di 521 byte, il router 1 crea due frame TCP per trasportare questo pacchetto al router peer DLSW 2. Il primo frame conterrà 520 byte di dati, 16 byte di informazioni DLSw e 40 byte di informazioni IP per un totale di 576 byte. Il secondo pacchetto conterrà 1 byte di dati, 16 byte di informazioni DLSw e 40 byte di informazioni IP. Quando si usa il PMTD e si presume che la MTU sia di 1500 byte per ottenere un valore MSS di 1460, il router 1 ha ricevuto dal router 2 un messaggio indicante che il router 2 può ricevere 1460 byte di dati. Il router 1 invia tutti i 521 byte di dati SDLC al router 2 in un pacchetto con 16 byte di informazioni DLSw e 40 byte di informazioni IP. Poiché DLSw è un evento di commutazione di processo, utilizzando PMTD, l'utilizzo della CPU per elaborare questo frame SDLC è stato dimezzato. Inoltre, il router 2 non deve attendere il secondo pacchetto per assemblare il frame LLC2. Con la funzionalità PMTD abilitata, il router 2 riceve l'intero pacchetto e può quindi rimuovere le informazioni IP e DLSw dal pacchetto e inviarlo allo switch 3745 senza ritardi.

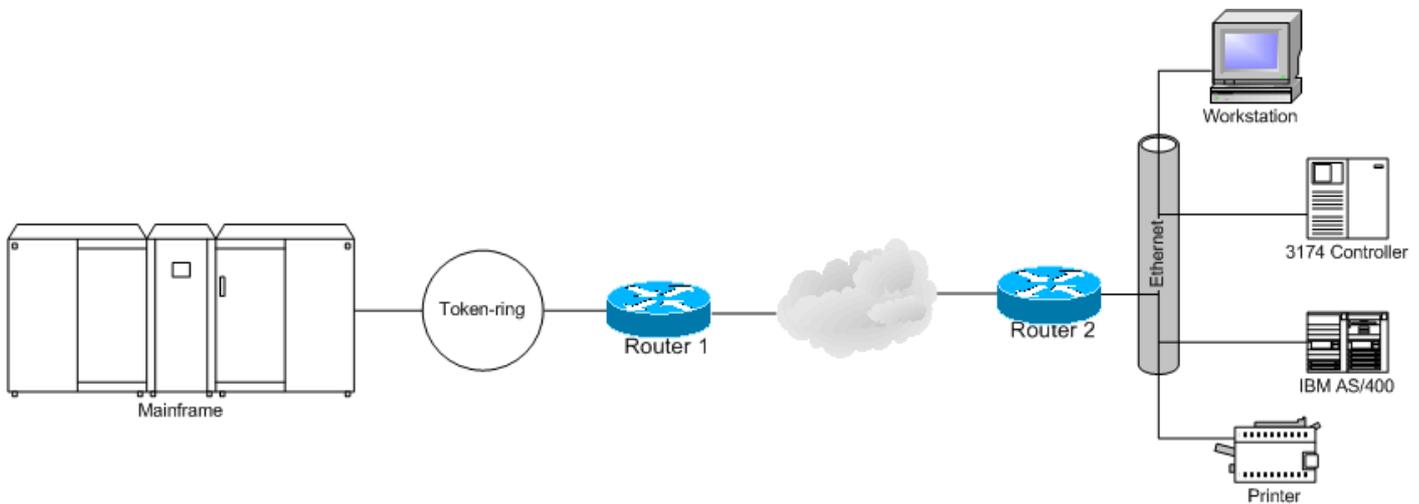
## Scenario 2 - Ritardo da pacchetti non ordinati



In questo scenario, sono disponibili due cloud IP con metriche uguali per il bilanciamento del carico o la ridondanza. La mancata attivazione di PMTD può rallentare notevolmente le DLSw. Senza PMTD abilitato, il router 1 deve assemblare il frame da 521 byte in due pacchetti TCP, uno con 520 byte di dati e il secondo con 1 byte di dati. Se il primo pacchetto attraversa il primo cloud IP, vi è una probabilità significativa che arrivi dopo il secondo pacchetto se il secondo pacchetto viene inviato tramite il cloud IP più basso e a costo uguale. In questo modo viene generato un pacchetto non ordinato. La capacità del protocollo IP/TCP è intrinseca alla capacità di gestire questo problema. I pacchetti non in ordine vengono archiviati in memoria in attesa dell'arrivo dell'intero flusso e quindi riassemblati. Non sono rari i pacchetti non ordinati; tuttavia, tutti i tentativi di ridurli a icona devono essere effettuati in quanto questo evento utilizza la memoria e le risorse della CPU. Una grande quantità di ordini non ordinati può causare un ritardo significativo a livello TCP. Se la sessione layer3/DLSw viene ritardata, la sessione LLC2/SDLC trasferita su questa DLSw ne risentirà. Se in questo scenario è abilitato il PMTD, viene inviato un singolo frame da 521 byte in un frame TCP su uno dei cloud IP. Il router ricevente richiede solo il buffer e decapsula un frame TCP.

La PMTD non ha alcuna relazione con la stazione terminale più grande annunciata dal frame in ambienti SNA. Ciò include il frame più grande (LF) nel campo RIF (Routing Information Field) su Token-Ring. La PMTD stabilisce in modo rigoroso la quantità di dati che possono essere incapsulati in un frame TCP. LLC2 e SDLC non hanno la capacità di frammentare i pacchetti, tuttavia, IP/TCP li ha. Un frame SNA di grandi dimensioni può essere segmentato perché è incapsulato in TCP. Questi dati vengono decapsulati sul router DLSw remoto e presentati nuovamente come dati SNA non frammentati.

## Scenario 3 - Connettività e throughput LLC2 più veloci



In questo scenario, lo switch 3174 e la workstation hanno sessioni attraverso il router TIC 3745 per il mainframe; se entrambi i dispositivi inviano dati destinati all'host, è possibile che il protocollo TCP possa inserire entrambi i frame LLC2 in un unico pacchetto. Senza PMTD, ciò non è possibile se i dati combinati dei due frame sono pari a 521 byte o superiori. In questo caso, il software TCP dovrà inviare ciascun pacchetto separatamente. Ad esempio, se lo switch 3174 e la workstation inviano un frame all'incirca contemporaneamente e ciascuno di questi pacchetti ha 400 byte di dati, il router riceve e memorizza ogni frame in un buffer. A questo punto, il router deve incapsulare ciascuno di questi flussi di dati da 400 byte in pacchetti TCP separati per la trasmissione al peer.

Con il PMTD abilitato e presupponendo un MSS di 1460, il router riceve e memorizza nel buffer i due pacchetti LLC2. e incapsularli entrambi in un unico pacchetto. Questo pacchetto TCP conterrà 40 byte di informazioni IP, 16 byte di informazioni DLSw per il primo accoppiamento di circuito DLSw, 400 byte di dati, altri 16 byte di dati per il secondo accoppiamento di circuito DLSw e gli altri 400 byte di dati. Questo scenario specifico utilizza due dispositivi e quindi due circuiti DLSw. PMTD consente alle DLSw di scalare in modo più efficiente un numero maggiore di circuiti DLSw. Molte reti spoke-hub richiedono centinaia di siti remoti, ciascuno con uno o due dispositivi SNA, che eseguono il peering in un router del sito centrale che si connette a un OSA o FEP fornendo accesso alle applicazioni host. La funzionalità PMTD consente ai protocolli TCP e DLSw di scalare in base a requisiti di dimensioni maggiori senza sovrautilizzare le risorse di memoria e CPU del router, oltre a fornire un trasporto più rapido.

**Nota:** un bug software è stato rilevato alla fine della versione 12.1(5)T e risolto nella versione 12.2(5)T, dove il PMTD non stava funzionando su un tunnel VPN (Virtual Private Network). Questo problema software è [CSCdt49552](#) (solo utenti [registrati](#)).

## [Verifica del PMTD per DLSW](#)

Eseguire il comando **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

Questa schermata viene identificata come sessione TCP DLSw perché una delle porte nella sessione TCP è 2065. Nella parte inferiore dell'output il segmento dati massimo è di 536 byte. Questo valore indica che sul router peer DLSw remoto della versione 10.1.1.1 il comando **ip tcp path-mtu-discovery** non è configurato. Il valore di 536 byte rappresenta già i 40 byte di informazioni IP in un frame IP. Questo valore di 536 byte non tiene conto dei 16 byte di informazioni DLSw che verrebbero aggiunte a un pacchetto TCP che trasporta il traffico SNA.

Con il comando **ip tcp path-mtu-discovery** configurato, il segmento di dati massimo è ora 1460. Inoltre, l'output del comando **show tcp** indica la **mtu** del **percorso** immediatamente precedente all'istruzione **max data segment**. L'interfaccia in uscita ha una MTU di 1500 byte. MTU è pari a 1500 byte meno 40 byte di informazioni IP sono 1460 byte. DLSw utilizzerà altri 16 byte. Pertanto, è possibile trasmettere fino a 1444 byte di frame LLC2 o SDLC in un frame 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

## [Informazioni correlate](#)

- [Note tecniche sui sistemi compatibili: Frammentazione IP e rilevamento del percorso MTU con VPN](#)
- [Supporto tecnico – Cisco Systems](#)