

IP 경로 MTU 검색 및 DLSw

목차

[소개](#)

[시작하기 전에](#)

[표기 규칙](#)

[사전 요구 사항](#)

[사용되는 구성 요소](#)

[배경 정보](#)

[PMTD를 사용하는 DLSw](#)

[DLSw에 대한 PMTD 확인](#)

[관련 정보](#)

소개

IBM 프로토콜 제품군, DLSw, STUN 및 BSTUN은 한 라우터에서 다른 라우터로 IP 세션 파이프를 설정합니다. TCP는 신뢰성 때문에 라우터 간 전송 방법으로 일반적으로 사용됩니다. 이 문서에서는 프래그먼트를 최소화하고 효율성을 극대화하는 세션 파이프에서 사용할 수 있는 가장 큰 MTU를 동적으로 검색하는 TCP의 기능에 대한 정보를 제공합니다.

시작하기 전에

표기 규칙

문서 규칙에 대한 자세한 내용은 [Cisco 기술 팁 표기 규칙](#)을 참조하십시오.

사전 요구 사항

이 문서에 대한 특정 요건이 없습니다.

사용되는 구성 요소

이 문서는 특정 소프트웨어 및 하드웨어 버전으로 한정되지 않습니다.

이 문서의 정보는 특정 랩 환경의 디바이스를 토대로 작성되었습니다. 이 문서에 사용된 모든 디바이스는 초기화된(기본) 컨피그레이션으로 시작되었습니다. 라이브 네트워크에서 작업하는 경우, 사용하기 전에 모든 명령의 잠재적인 영향을 이해해야 합니다.

배경 정보

RFC 1191에 설명된 대로 PMTD(Path MTU Discovery)는 IP 패킷의 기본 바이트 크기가 576임을

지정합니다. 프레임의 IP 및 TCP 부분은 40바이트에서 536바이트를 데이터 페이로드로 남깁니다. 이 공간을 최대 세그먼트 크기 또는 MSS라고 합니다. RFC1191의 섹션 3.1은 협상할 수 있는 더 큰 MSS를 지정하며, 이것이 바로 Cisco 라우터에서 **ip tcp path-mtu-discovery** 명령을 실행하는 방법입니다. 이 명령이 구성되고 TCP 세션이 시작되면 라우터의 SYN 패킷에 더 큰 MSS를 지정하는 TCP 옵션이 포함됩니다. 이 더 큰 MSS는 아웃바운드 인터페이스의 MTU에서 40바이트를 뺀 것입니다. 아웃바운드 인터페이스의 MTU가 1500바이트이면 알려진 MSS는 1460입니다. 아웃바운드 인터페이스에 더 큰 MTU가 있는 경우(예: 4096바이트 MTU가 포함된 프레임 릴레이) MSS는 4096바이트에서 40바이트의 IP 정보를 뺀 값으로 **show tcp** 명령 출력(최대 데이터 세그먼트는 4056바이트)에 표시됩니다.

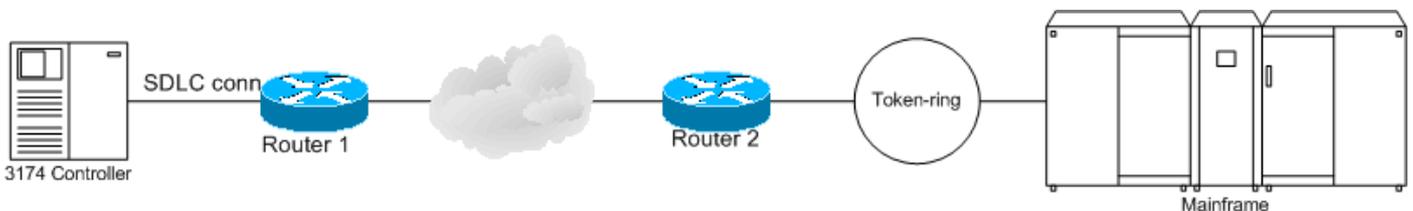
라우터에서 PMTD를 구성해도 라우터에서/라우터로 이미 설정된 기존 TCP 세션에는 영향을 미치지 않습니다. PMTD는 11.3.5T IOS 레벨에 소개되었으며, IOS의 후속 릴리스에서 선택적 명령이 되었습니다. IOS 11.3(5)T 이전에는 기본적으로 켜져 있었습니다. 또한 PMTD는 IP 주소가 동일한 서브넷에 있는 경우 자동으로 이루어지며, 이는 IP 주소가 동일한 미디어에 직접 연결되어 있음을 나타냅니다.

PMTD가 제대로 작동하려면 두 라우터를 모두 구성해야 합니다. 두 라우터가 모두 구성되면 한 라우터에서 다른 라우터로 SYN에는 더 높은 MSS를 광고하는 선택적 TCP 값이 포함됩니다. 그러면 반환 SYN이 더 높은 MSS 값을 알립니다. 따라서 양쪽은 더 큰 MSS를 받아들일 수 있다고 다른 측에 알립니다. 라우터 1에 **ip tcp path-mtu-discovery** 명령이 하나만 구성되어 있으면 더 큰 MSS를 광고하므로 라우터 2는 라우터 1에 1460바이트 프레임을 전송할 수 있습니다. 라우터 2는 더 큰 MSS를 광고하지 않으므로 라우터 1은 기본값을 전송하도록 잠깁니다.

PMTD를 사용하는 DLSw

IBM 제품군에서 DLSw, STUN 및 BSTUN은 라우터에서 라우터로 TCP 세션을 통해 대량의 데이터를 전달하는 임무를 수행할 수 있습니다. PMTD는 11.2 및 이전 IOS 레벨에서 기본적으로 활성화되었음을 고려하여 PMTD를 구현하는 것이 매우 중요하고 유익할 수 있습니다. RFC에 따라 가장 큰 프레임은 기본적으로 576바이트이며 IP/TCP 캡슐화의 경우 40바이트입니다. DLSw는 캡슐화에 다른 16바이트를 사용합니다. 기본 MSS를 사용하여 전송 중인 실제 데이터는 520바이트입니다. 또한 DLSw는 두 개의 서로 다른 LLC2(Logical Link Control 2) 패킷을 하나의 TCP 프레임으로 전달할 수 있습니다. 두 워크스테이션이 각각 LLC2 프레임을 전송하는 경우 DLSw는 두 LLC2 프레임을 하나의 프레임에서 DLSw 원격 피어로 전달할 수 있습니다. MSS가 클수록 TCP 드라이버는 이 piggybacking 스키마를 수용할 수 있습니다. 다음은 path-mtu-discovery 명령의 값을 설명하기 위한 세 가지 주요 시나리오입니다.

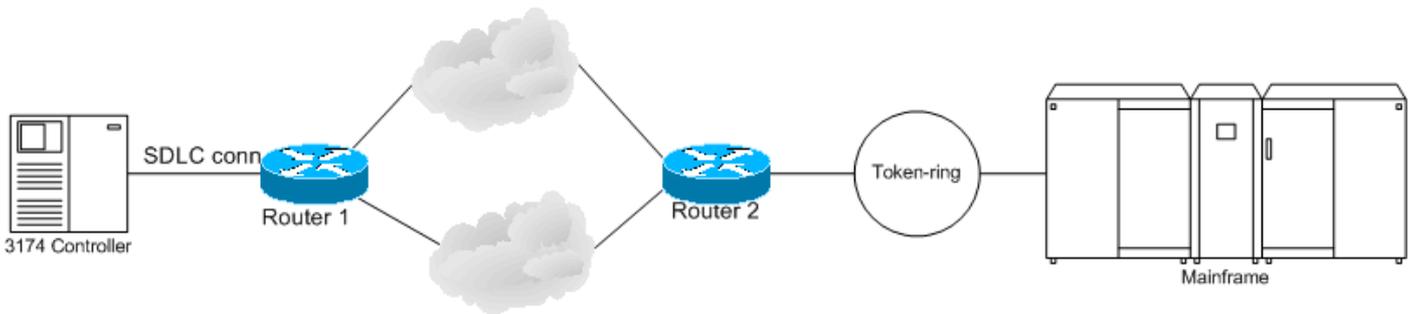
시나리오 1 - 원치 않는 오버헤드



일반적으로 SDLC 디바이스는 각 프레임에서 최대 데이터 265 또는 521바이트에 대해 구성됩니다. 값이 521이고 3174가 라우터 1에 521바이트 SDLC 프레임으로 보내는 경우 라우터 1은 이를 DLSW 피어 라우터 2로 전송하기 위해 두 개의 TCP 프레임을 만듭니다. 첫 번째 프레임에는 520바이트의 데이터, 16바이트의 DLSw 정보, 40바이트의 IP 정보가 포함되며 총 576바이트입니다. 두 번째 패킷에는 1바이트의 데이터, 16바이트의 DLSw 정보, 40바이트의 IP 정보가 포함됩니다. PMTD를 사용하고 1460 MSS를 얻기 위해 1500바이트 MTU를 가정할 때 라우터 2에서 라우터 1에 라우터 2가 1,460바이트의 데이터를 수신할 수 있다고 들었습니다. 라우터 1은 모든 521바이트의 SDLC 데이터를 16바이트의 DLSw 정보와 40바이트의 IP 정보가 포함된 하나의 패킷으로 라우

터 2에 전송합니다.DLSw는 프로세스 전환 이벤트이므로 PMTD를 사용하여 이 SDLC 프레임을 처리하는 CPU 사용률이 절반으로 감소했습니다.또한 라우터 2는 두 번째 패킷이 LLC2 프레임을 어셈블할 때까지 기다릴 필요가 없습니다.PMTD가 활성화된 경우 라우터 2는 전체 패킷을 수신한 다음 패킷에서 IP 및 DLSW 정보를 제거한 다음 이를 3745로 즉시 전송할 수 있습니다.

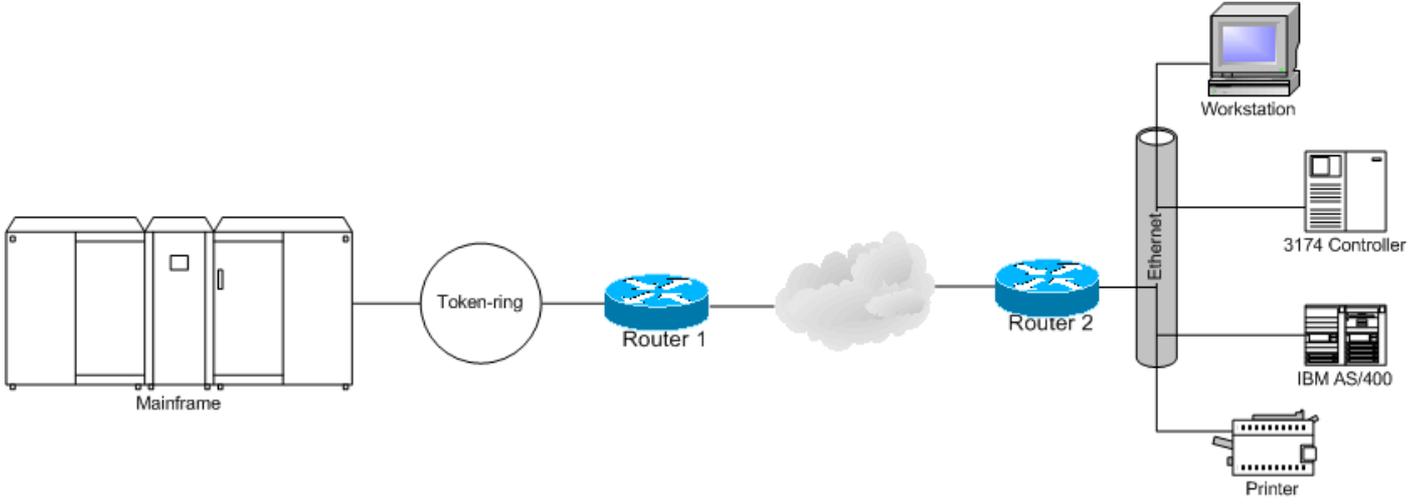
시나리오 2 - Out-of-Order 패킷에서 지연



이 시나리오에서는 로드 밸런싱 또는 리던던시를 위해 동일한 메트릭으로 두 개의 IP 클라우드를 사용할 수 있습니다.PMTD를 활성화하지 않으면 DLSw가 크게 느려질 수 있습니다.PMTD를 활성화하지 않으면 라우터 1은 521바이트 프레임을 2개의 TCP 패킷으로 어셈블해야 합니다. 1은 520바이트의 데이터를, 2번째는 1바이트의 데이터를 포함해야 합니다.첫 번째 패킷이 상위 IP 클라우드를 통과하는 경우, 두 번째 패킷이 더 낮은 동일 비용 IP 클라우드를 통해 전송될 경우 두 번째 패킷 이후에 도달하게 될 가능성이 높습니다.이렇게 하면 순서가 잘못된 패킷이라고 하는 것이 생성됩니다.IP/TCP 프로토콜의 기능에는 이 문제를 관리하는 기능이 있습니다.순서가 잘못된 패킷은 전체 스트림이 도착하고 다시 어셈블될 때까지 기다리는 메모리에 저장됩니다.순서가 잘못된 패킷은 일반적이지 않지만, 이 이벤트가 메모리 및 CPU 리소스를 사용하므로 이러한 패킷을 최소화하려는 모든 시도가 이루어져야 합니다.많은 양의 주문량이 초과 처리되면 TCP 레벨에서 상당한 지연이 발생할 수 있습니다.레이어3/DLSw 세션이 지연되면 이 DLSw를 통해 전송되는 LLC2/SDLC 세션이 이후에 손상됩니다.이 시나리오에서 PMTD가 활성화된 경우 단일 521바이트 프레임이 IP 클라우드를 통해 하나의 TCP 프레임으로 전송됩니다.수신 라우터에는 버퍼만 필요하며 하나의 TCP 프레임을 캡슐화해야 합니다.

PMTD는 SNA 환경에서 가장 큰 프레임 광고된 엔드 스테이션과 엔드 스테이션과의 관계가 없습니다.여기에는 토큰 링의 RIF(Routing Information Field)의 LF(Largest Frame)가 포함됩니다.PMTD는 하나의 TCP 프레임으로 캡슐화할 수 있는 데이터의 양을 엄격하게 지정합니다.LLC2 및 SDLC에는 기능 프래그먼트 패킷이 없지만 IP/TCP에는 없습니다.큰 SNA 프레임은 TCP로 캡슐화되므로 세그멘테이션할 수 있습니다.이 데이터는 원격 DLSw 라우터에서 역캡슐화되고 다시 조각화되지 않은 SNA 데이터로 표시됩니다.

시나리오 3 - 더 빠른 LLC2 연결 및 처리량



이 시나리오에서 3174와 워크스테이션은 3745 TIC을 통해 메인프레임으로 세션을 제공합니다. 두 디바이스 모두 호스트로 향하는 데이터를 전송하면 TCP에서 두 LLC2 프레임을 하나의 패킷에 넣을 수 있습니다. PMTD가 없으면 두 프레임의 결합된 데이터가 521바이트 이상인 경우 이 작업을 수행할 수 없습니다. 이 경우 TCP 소프트웨어는 각 패킷을 별도로 전송해야 합니다. 예를 들어 3174와 워크스테이션이 거의 동시에 프레임을 전송하고 이러한 각 패킷에 400바이트의 데이터가 있는 경우 라우터는 각 프레임을 수신하고 버퍼링합니다. 이제 라우터는 이러한 400바이트 데이터 스트림을 각각 별도의 TCP 패킷으로 캡슐화하여 피어로 전송해야 합니다.

PMTD가 활성화되고 MSS가 1460이라고 가정하면 라우터는 2개의 LLC2 패킷을 수신하고 버퍼링합니다. 이제 둘 다 하나의 패킷으로 캡슐화할 수 있습니다. 이 하나의 TCP 패킷에는 40바이트의 IP 정보, 첫 번째 DLSw 회선 페어링을 위한 16바이트의 DLSw 정보, 400바이트의 데이터, 두 번째 DLSw 회선 페어링을 위한 또 다른 16바이트의 데이터 및 다른 400바이트의 데이터가 포함됩니다. 이 특정 시나리오에서는 두 개의 디바이스, 즉 두 개의 DLSw 회로를 사용합니다. PMTD를 통해 DLSw는 더 많은 수의 DLSw 회선으로 더 효율적으로 확장할 수 있습니다. 많은 스포크 허브 네트워크에는 각각 1~2개의 SNA 디바이스가 있는 수백 개의 원격 사이트가 필요하며, OSA 또는 FEX에 연결하는 중앙 사이트 라우터로 피어링하여 호스트 애플리케이션에 대한 액세스를 제공합니다. PMTD는 라우터 CPU와 메모리 리소스를 과다 사용하지 않고 빠른 전송을 제공하지 않고도 TCP 및 DLSw에서 더 큰 요구 사항으로 확장할 수 있는 기능을 제공합니다.

참고: 12.1(5)T 후반에 발견된 소프트웨어 버그가 12.2(5)T에서 해결되었으며 PMTD가 VPN(Virtual Private Network) 터널을 통해 작동하지 않았습니다. 이 소프트웨어 결함은 [CSCdt49552입니다](#)(등록된 고객만 해당).

DLSW에 대한 PMTD 확인

show tcp 명령을 실행합니다.

```
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	Wakeup	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

이 표시는 TCP 세션의 포트 중 하나가 2065이므로 DLSw TCP 세션으로 식별됩니다.출력 하단의 최대 데이터 세그먼트는 536바이트입니다.이 값은 10.1.1.1의 원격 DLSw 피어 라우터에 ip tcp path-mtu-discovery 명령이 구성되지 않았음을 나타냅니다.536바이트 값은 IP 프레임의 IP 정보 40바이트를 이미 계산합니다.이 536바이트 값은 SNA 트래픽을 전달하는 TCP 패킷에 추가될 16바이트의 DLSw 정보를 고려하지 않습니다.

ip tcp path-mtu-discovery 명령이 구성된 경우 최대 데이터 세그먼트는 이제 1460입니다.또한 show tcp 명령 출력은 max data segment 문 바로 앞에 경로 mtu를 나타냅니다.아웃바운드 인터페이스의 MTU는 1500바이트입니다.MTU는 1500바이트에서 40바이트의 IP 정보를 뺀 1460바이트입니다 .DLSw는 다른 16바이트를 사용합니다.따라서 LLC2 또는 SDLC의 최대 1444바이트 프레임을 하나의 TCP 프레임으로 전송할 수 있습니다.

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

관련 정보

- [호환 시스템 기술 노트:VPN을 통한 IP 조각화 및 MTU 경로 검색](#)
- [Technical Support - Cisco Systems](#)