

Cisco IOS XRおよびBGPでのパスMTUディスカバリの確認

内容

[概要](#)

[背景説明](#)

[TCP PMTUDおよびTCP MSS](#)

[シナリオ：TCP PMTUDが無効](#)

[デフォルトのMTU値の使用](#)

[デフォルト以外のMTU値の使用 – アクティブTCPピア](#)

[デフォルト以外のMTU値の使用：パッシブTCPピア](#)

[TCPオプションの使用 – XRアクティブ](#)

[TCPオプションの使用：XRパッシブ](#)

[直接接続されていないTCPピア](#)

[直接接続されていないTCPピア – TCPオプションの使用\(MD5\)](#)

[直接接続されていないTCPピア：パスセグメントのIP MTUが小さい](#)

[シナリオ：TCP PMTUD有効](#)

[PMTUDの有効化](#)

[PMTUD：パスセグメントのIP MTUが小さい](#)

[PMTUD:TCPオプション\(MD5\)](#)

[PMTUD – ブラックホール検出](#)

概要

このドキュメントでは、Cisco IOS® XRデバイスのTransmission Control Protocol(TCP)Path Maximum Transmission Unit(MTU)Discovery(PMTUD)について説明します。

背景説明

PMTUDメカニズムは、2つのホスト間のパス上のどこでもフラグメンテーションを必要としない最大のインターネットプロトコル(IP)パケットサイズを決定しようとします。確立された値はパスMTUと指定され、各ホップのMTU値の最小値と等しくなります。情報を送信するときにパスMTUを考慮すると、ネットワーク容量を最大限に活用し、フラグメンテーションと伝送効率を回避できます。PMTUDの仕組みと実装は、PMTUDの動作を徐々に明らかにするクライアントプロトコルとしてボーダーゲートウェイプロトコル(BGP)を使用して、さまざまなシナリオにわたって導入されます。

TCP PMTUDおよびTCP MSS

TCPは、ローカルの最大セグメントサイズ(MSS)に影響を与えるためにPMTUDの結果を活用します。つまり、検出されたパスMTUに動的に適応します。したがって、PMTUDに移行する前に、TCP最大セグメントサイズ(MSS)を迅速に確認し、その意味とその目的を理解できます。

RFC879のMSSの元の定義に[従って](#):MSSオプションの定義は、次のように記述できます。IPヘッダーオプションのないIPデータグラムで送信されるTCPヘッダーオプションのないTCPセグメントで、このTCPオプションの送信者が受信できるデータオクテットの最大数。

いくつかの側面を明らかにし、実装者にアドバイスを提供するため、[RFC6691](#) は、MSS値の計算方法を示しています。

TCP MSSオプションに含める値を計算する場合、MTU値は固定IPおよびTCPヘッダーのサイズだけ小さくし、IPまたはTCPオプションを考慮して小さくしないでください。逆に、送信側は、送信するパケットに含まれるIPまたはTCPオプションを考慮して、TCPデータ長を短くする必要があります。

MSSの詳細な定義は、『[Routing Configuration Guide for Cisco ASR 9000 Series Routers, IOS XR Release 6.7.x](#)』から抽出できます。

MSSは、1つのフラグメント化されていないTCPセグメントでコンピュータまたは通信デバイスが受信できる最大のデータ量です。すべてのTCPセッションは、1つのパケットで転送できるバイト数の制限によって制限されます。この制限はMSSです。TCPは、パケットをIP層に渡す前に、送信キュー内のチャンクにパケットを分割します。

TCP MSS値は、インターフェイスのMTUに依存します。MTUは、1つのインスタンスでプロトコルによって送信できるデータの最大長です。最大TCPパケット長は、送信元デバイスの発信インターフェイスのMTUと、TCPセットアッププロセス中に宛先デバイスによってアナウンスされたMSSの両方によって決定されます。MSSがMTUに近いほど、BGPメッセージの転送がより効率的になります。データフローの各方向は、異なるMSS値を使用できます。

特定のTCPセッションのMSSについてTCPが考慮すべき値は何ですか。どのように計算されるのでしょうか？

[RFC879](#)に基づくデフォルト値には、[次の項目](#)があります。ホストは、宛先ホストが大きなデータグラムを受け入れる準備ができていないことを特定の知識がない限り、576オクテットを超えるデータグラムを送信できません。TCP最大セグメントサイズはIP最大データグラムサイズから40を引いた値です。

デフォルトのIPデータグラムの最大サイズは576です。

デフォルトのTCP最大セグメントサイズは536です。

これは、576バイトのIP MTU値を考慮します。ただし、実際のIP MTU値を無視すると、TCP MSSの計算は次のように要約できます。

- アクティブピア：SYNパケットを使用して初期MSSを計算し、送信します。

$MSS = IPMTU - \text{sizeof}(\text{minimum TCPHDR}) - \text{sizeof}(\text{minimum IPHDR})$

Where,

$\text{sizeof}(\text{minimum TCPHDR}) = 20 \text{ bytes.}$

$\text{sizeof}(\text{minimum IPHDR}) = 20 \text{ bytes.}$

- Passive Peer：初期MSSを計算し、アクティブピアから受信したMSSと比較し、SYN、ACKを、これらのMSS値の低い方で送信します。

$\text{MIN}[\text{IPMTU} - \text{sizeof}(\text{minimum TCPHDR}) - \text{sizeof}(\text{minimum IPHDR}), \text{Received MSS value}]$

Where,

$\text{sizeof}(\text{minimum TCPHDR}) = 20 \text{ bytes.}$

$\text{sizeof}(\text{minimum IPHDR}) = 20 \text{ bytes.}$

Received MSS value = MSS value received with Active Peer TCP SYN.

MSSオプション値に関するネゴシエーションはありません。各ノードは独自の値を決定し、TCPセッションが確立されると同じ値をアナウンスします。MSS計算で考慮されるIP MTU値がPMTUDから取得できる場合、MSS値を特定のパスMTUの最も有効な値に調整できることが明らかになります。Cisco IOS XRの動作には、MSSの計算とPMTUDの役割に関するいくつかの詳細があります。

PMTUDは、Cisco IOS XRではデフォルトで無効になっています。

- ローカルの初期MSS計算では、次のようにIP MTUが考慮されます。直接接続されたピアの場合：出インターフェイスのIP MTUを考慮します。直接接続されていないピアの場合：1280バイトのIP MTUを考慮します。MSS値は、設定されたTCPオプションの影響を受けます。

Cisco IOS XRでPMTUDが有効になっている場合：

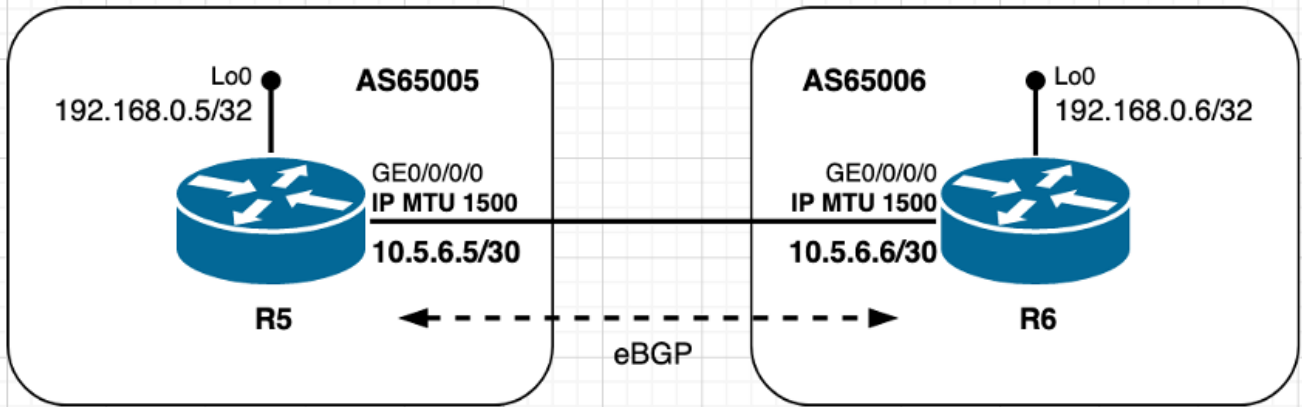
- ローカルの初期MSS計算では、次のようにIP MTUが考慮されます。直接接続されたピアまたは直接接続されていないピアに関係なく、出インターフェイスのIP MTUを考慮します。MSS値は、設定されたTCPオプションの影響を受けます。

PMTUDの仕組みと実装に関する詳細については、考慮する必要があり、次の表に示す実用的な例を通じて、このドキュメントで紹介されている内容です。この表では、アクティブおよびパッシブTCPピアのIP MTUと、考慮される各シナリオの選択されたMSS値も示します。

PMTUD	Scenarios	ACTIVE IP MTU	PASSIVE IP MTU	MSS
Disabled	Using default MTU values	1500	1500	1460
	Using non-default MTU value – Active TCP peer	4460	1500	1460
	Using non-default MTU value – Passive TCP peer	1500	4460	1460
	Using TCP Options (MD5) – XR Active	1500	1500	1436
	Using TCP Options (MD5) – XR Passive	1500	1500	1460
	TCP peers not directly connected	1500	1500	1240
	TCP peers not directly connected – Using TCP Options (MD5)	1500	1500	1216
Enabled	Enabling TCP PMTUD	1500	1500	1460
	PMTUD in action – Path segment has lower MTU	1500	1500	1460
	PMTUD in action – TCP Options (MD5)	1500	1500	1436

シナリオ：TCP PMTUDが無効

デフォルトのMTU値の使用



イメージ2.1.デフォルトのMTU値を使用

図2.1 R6に示すeBGPピアがTCP接続を管理する場合、アクティブな役割を果たし、宛先ポート179でR5とのTCPセッションを開始します。ピアは直接接続され、両方とも各インターフェイスのデフォルトIP MTU値を使用します。このドキュメントの冒頭で共有されている情報に基づいて、このシナリオのMSS計算を次のように要約できます。

- 両方のノードが1500バイトのデフォルトIP MTUを使用します
- TCPパスMTUディスカバリはデフォルトで無効になっています
- TCPピアが直接接続されている R6はBGP接続を管理しますR6は1460バイトのMSSでSYNを送信します 1500 (インターフェイスIP MTU) ~ 20(minTCP_H) ~ 20(minIP_H)R5は1460バイトのMSSでSYN、ACKを送信します [Received MSS;ローカル初期MSS]受信MSS 1460バイト。ローカル初期MSS 1460バイト最小のMSS値が両方のピアで使用されます

R6に表示されるTCPセッションの詳細 – ACTIVE:

! - As seen on R6 - ACTIVE

```
RP/0/0/CPU0:R6#show interfaces gigabitEthernet 0/0/0/0
Fri Jan 8 09:35:48.553 UTC
GigabitEthernet0/0/0/0 is up, line protocol is up
Interface state transitions: 1
Hardware is GigabitEthernet, address is fa16.3e85.3dc2 (bia fa16.3e85.3dc2)
Internet address is 10.5.6.6/30
MTU 1514 bytes, BW 1000000 Kbit (Max: 1000000 Kbit)
<snip>
```

```
RP/0/0/CPU0:R6#show tcp brief
Fri Jan 8 09:36:22.491 UTC
PCB      VRF-ID      Recv-Q  Send-Q  Local Address          Foreign Address        State
<snip>
0x121649fc 0x60000000      0       0      10.5.6.6:24454        10.5.6.5:179          ESTAB
<snip>
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x121649fc
Fri Jan 8 09:37:00.888 UTC
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 09:28:28 2021
```

```
PCB 0x121649fc, SO 0x121561b8, TCPCB 0x12156f64, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 78
Local host: 10.5.6.6, Local port: 24454 (Local App PID: 1011918)
```

Foreign host: 10.5.6.5, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	13	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	10	2	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 3757770712 snduna: 3757770960 sndnxt: 3757770960

sndmax: 3757770960 sndwnd: 32574 sndcwnd: 4380

irs: 1072103647 rcvnxt: 1072103895 rcvwnd: 32593 rcvadv: 1072136488

SRTT: 155 ms, RTTO: 540 ms, RTV: 385 ms, KRTT: 0 ms

minRTT: 9 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec

Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE

Connect retries remaining: 30, connect retry interval: 50 secs

State flags: none

Feature flags: Win Scale, Nagle

Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0

Timestamp option: recent 0, recent age 0, last ACK sent 0

Sack blocks {start, end}: none

Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO

Socket states: SS_ISCONNECTED, SS_PRIV

Socket receive buffer states: SB_DEL_WAKEUP

Socket send buffer states: SB_DEL_WAKEUP

Socket receive buffer: Low/High watermark 1/32768

Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:

#PDU's in buffer: 0

FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:

Num Labels: 0 Label Stack:

RP/0/0/CPU0:R6

R5に表示されるTCPセッションの詳細: パッシブ:

! - As seen on R5 - PASSIVE

RP/0/0/CPU0:R5#show interfaces gigabitEthernet 0/0/0/0

Fri Jan 8 09:33:04.564 UTC

GigabitEthernet0/0/0/0 is up, line protocol is up

Interface state transitions: 1

Hardware is GigabitEthernet, address is fa16.3ead.518f (bia fa16.3ead.518f)

Internet address is 10.5.6.5/30

MTU 1514 bytes, BW 1000000 Kbit (Max: 1000000 Kbit)

<snip>

RP/0/0/CPU0:R5#show tcp brief

Fri Jan 8 09:33:53.221 UTC

PCB	VRF-ID	Recv-Q	Send-Q	Local Address	Foreign Address	State
<snip>						
0x12155884	0x60000000	0	0	10.5.6.5:179	10.5.6.6:24454	ESTAB

<snip>

RP/0/0/CPU0:R5#show tcp detail pcb 0x12155884

Fri Jan 8 09:34:47.317 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 09:28:29 2021

PCB 0x12155884, SO 0x1215568c, TCPCB 0x12155a54, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 78
Local host: 10.5.6.5, Local port: 179 (Local App PID: 1044686)
Foreign host: 10.5.6.6, Foreign port: 24454

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	9	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	9	7	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 1072103647 snduna: 1072103857 sndnxt: 1072103857
sndmax: 1072103857 sndwnd: 32631 sndcwnd: 4380
irs: 3757770712 rcvnxt: 3757770922 rcvwnd: 32612 rcvadv: 3757803534

SRTT: 47 ms, RTTO: 300 ms, RTV: 170 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 219 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

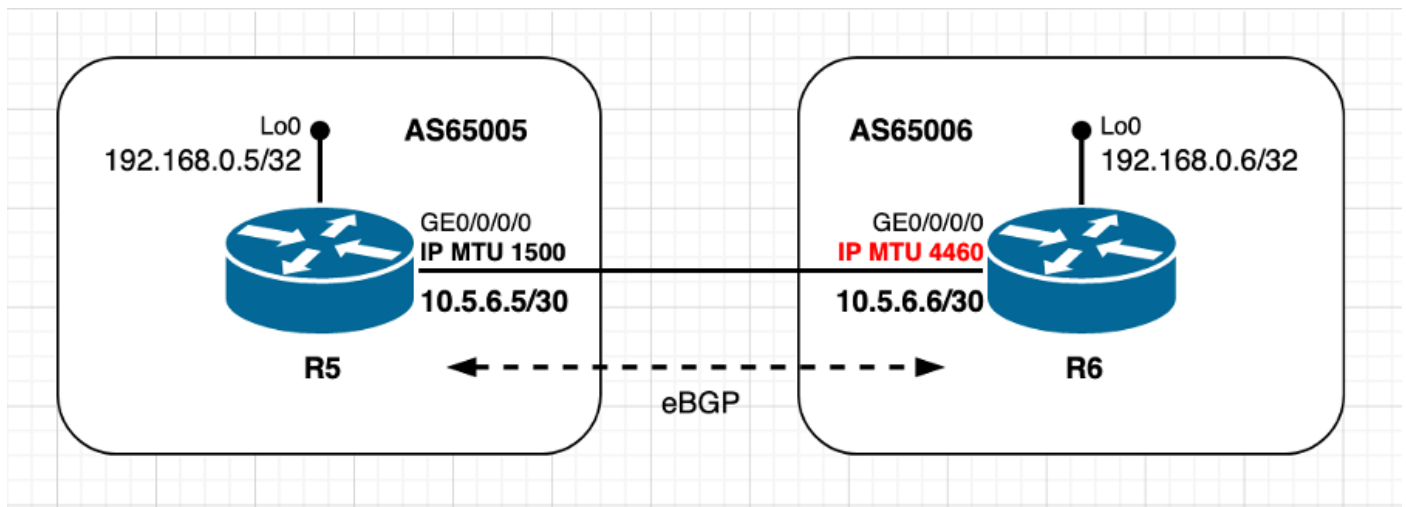
Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:

```
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:
```

```
RP/0/0/CPU0:R5#
```

デフォルト以外のMTU値の使用 – アクティブTCPピア



イメージ2.2:ACTIVEピアはデフォルト以外のMTU値を使用する

このシナリオは基本的に前と同じですが、アクティブなTCPピアR6がデフォルト以外のIP MTU値を使用している唯一の違いがあります。MSS値の初期計算と決定がパッシブTCPピアR5によってどのように行われるかに注意してください。このシナリオのTCP MSS計算は、次のように要約できます。

- R6はデフォルト以外のIP MTU 4460バイトを使用します
- TCPパスMTUディスカバリーはデフォルトで無効になっています
- TCPピアが直接接続されている R6はBGP接続を管理しますR6はSYNを4420バイトのMSSで送信 4460 (インターフェイスIP MTU) ~ 20(minTCP_H) ~ 20(minIP_H)R5はSYNを送信し、MSSが1460バイトのACK [Received MSS;ローカル初期MSS]受信MSS 4420バイト。ローカル初期MSS 1460バイト最小のMSS値が両方のピアで使用されます

R6から送信されたTCP SYN:

```
! - TCP SYN sourced from R6
```

```
140    1598.150521    10.5.6.6        10.5.6.5        TCP    62      35502  179 [SYN] Seq=0
Win=16384 Len=0  MSS=4420 WS=1
```

```
Frame 140: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:85:3d:c2 (fa:16:3e:85:3d:c2), Dst: fa:16:3e:ad:51:8f
(fa:16:3e:ad:51:8f)
```

```
Internet Protocol Version 4, Src: 10.5.6.6, Dst: 10.5.6.5
```

```
Transmission Control Protocol, Src Port: 35502, Dst Port: 179, Seq: 0, Len: 0
```

```
Source Port: 35502
```

```
Destination Port: 179
```

```
[Stream index: 6]
```

```
[TCP Segment Len: 0]
```

```
Sequence number: 0 (relative sequence number)
```

```
Acknowledgment number: 0
```

```
Header Length: 28 bytes
```

```
Flags: 0x002 (SYN)
```

```
Window size value: 16384
```

```
[Calculated window size: 16384]
Checksum: 0x219d [unverified]
[Checksum Status: Unverified]
Urgent pointer: 0
Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
  Maximum segment size: 4420 bytes
    Kind: Maximum Segment Size (2)
    Length: 4
    MSS Value: 4420
  Window scale: 0 (multiply by 1)
  End of Option List (EOL)
```

TCP SYN、R5からACK:

! - TCP SYN, ACK sourced from R5

```
141    1598.154866    10.5.6.5        10.5.6.6        TCP        62        179    35502 [SYN, ACK] Seq=0
Ack=1 Win=16384 Len=0 MSS=1460 WS=1
```

```
Frame 141: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:ad:51:8f (fa:16:3e:ad:51:8f), Dst: fa:16:3e:85:3d:c2
(fa:16:3e:85:3d:c2)
Internet Protocol Version 4, Src: 10.5.6.5, Dst: 10.5.6.6
Transmission Control Protocol, Src Port: 179, Dst Port: 35502, Seq: 0, Ack: 1, Len: 0
  Source Port: 179
  Destination Port: 35502
  [Stream index: 6]
  [TCP Segment Len: 0]
  Sequence number: 0      (relative sequence number)
  Acknowledgment number: 1    (relative ack number)
  Header Length: 28 bytes
  Flags: 0x012 (SYN, ACK)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0xe2b4 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
    Maximum segment size: 1460 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1460
    Window scale: 0 (multiply by 1)
    End of Option List (EOL)
```

R6に表示されるTCPセッションの詳細 - ACTIVE:

! - as seen on R6 - Active

```
RP/0/0/CPU0:R6#show interfaces gigabitEthernet 0/0/0/0
Fri Jan  8 09:46:54.138 UTC
GigabitEthernet0/0/0/0 is up, line protocol is up
  Interface state transitions: 1
  Hardware is GigabitEthernet, address is fa16.3e85.3dc2 (bia fa16.3e85.3dc2)
  Internet address is 10.5.6.6/30
  MTU 4474 bytes, BW 1000000 Kbit (Max: 1000000 Kbit)
<snip>
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1215761c
Fri Jan  8 09:56:25.819 UTC
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
```


Established at Fri Jan 8 09:51:46 2021

PCB 0x1215761c, SO 0x12156f64, TCPCB 0x1216419c, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 886
Local host: 10.5.6.6, Local port: 35502 (Local App PID: 1011918)
Foreign host: 10.5.6.5, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	9	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	6	5	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 764231407 snduna: 764231579 sndnxt: 764231579
sndmax: 764231579 sndwnd: 32650 sndcwnd: 4380
irs: 2712512697 rcvnxt: 2712512869 rcvwnd: 32669 rcvadv: 2712545538

SRTT: 31 ms, RTTO: 300 ms, RTV: 130 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 50 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 4420, max MSS 4420

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R6#

R5に表示されるTCPセッションの詳細: パッシブ:

! - as seen on R5 - Passive

RP/0/0/CPU0:R5#show tcp detail pcb 0x12155a98
Fri Jan 8 09:55:18.193 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 09:51:47 2021

PCB 0x12155a98, SO 0x12153ea0, TCPCB 0x12154e18, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 886
Local host: 10.5.6.5, Local port: 179 (Local App PID: 1044686)
Foreign host: 10.5.6.6, Foreign port: 35502

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	6	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	6	1	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 2712512697 snduna: 2712512850 sndnxt: 2712512850
sndmax: 2712512850 sndwnd: 32688 sndcwnd: 4380
irs: 764231407 rcvnxt: 764231560 rcvwnd: 32669 rcvadv: 764264229

SRTT: 107 ms, RTTO: 538 ms, RTV: 431 ms, KRTT: 0 ms
minRTT: 29 ms, maxRTT: 219 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

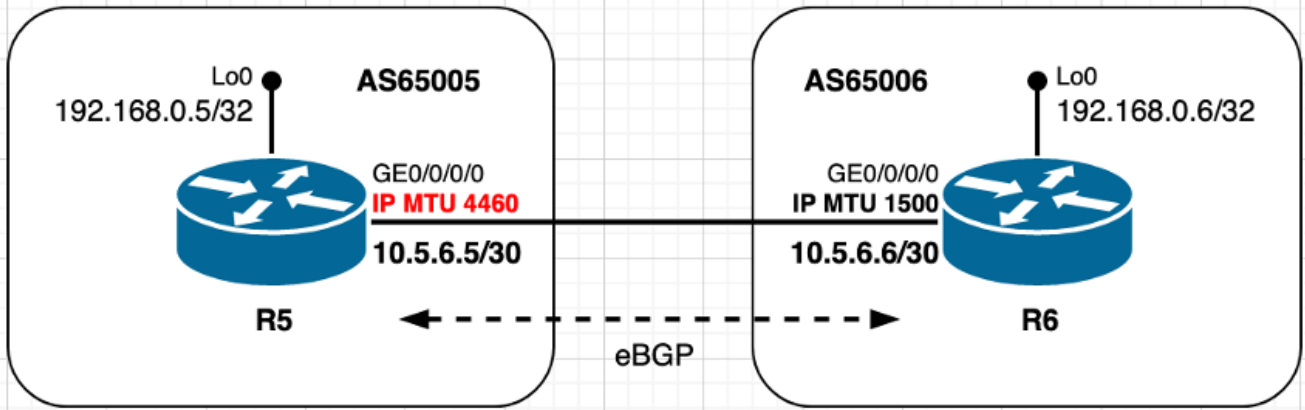
Datagrams (in bytes): MSS 1460, peer MSS 4420, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none
Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R5#

デフォルト以外のMTU値の使用 : パッシブTCPピア



イメージ2.3 : パッシブピアはデフォルト以外のMTU値を使用します。

それでも同じeBGPシナリオですが、パッシブTCPピアR5がデフォルト以外のIP MTUで設定され、アクティブTCPピアR6がデフォルトのIP MTU値で設定されています。前のシナリオと同様に、パッシブピアR5によってMSS値が選択される方法をメモします。このシナリオでのTCP MSSの計算は、次のように要約できます。

- R5はデフォルト以外のIP MTU 4460バイトを使用します
- TCPパスMTUディスカバリーはデフォルトで無効になっています
- TCPピアが直接接続されている R6はBGP接続を管理しますR6は1460バイトのMSSでSYNを送信します 1500 (インターフェイスIP MTU) ~ 20(minTCP_H) ~ 20(minIP_H)R5はSYNを送信し、MSSが1460バイトのACK [Received MSS;ローカル初期MSS]受信MSS 1460バイト。ローカル初期MSS 4420バイト最小のMSS値が両方のピアで使用されます

R6から送信されたTCP SYN:

! - TCP SYN sourced from R6

```
237      2696.666481      10.5.6.6      10.5.6.5      TCP      62      47007 179 [SYN] Seq=0
Win=16384 Len=0 MSS=1460 WS=1
```

```
Frame 237: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:85:3d:c2 (fa:16:3e:85:3d:c2), Dst: fa:16:3e:ad:51:8f
(fa:16:3e:ad:51:8f)
```

```
Internet Protocol Version 4, Src: 10.5.6.6, Dst: 10.5.6.5
```

```
Transmission Control Protocol, Src Port: 47007, Dst Port: 179, Seq: 0, Len: 0
```

```
Source Port: 47007
```

```
Destination Port: 179
```

```
[Stream index: 10]
```

```
[TCP Segment Len: 0]
```

```
Sequence number: 0 (relative sequence number)
```

```
Acknowledgment number: 0
```

```
Header Length: 28 bytes
```

```
Flags: 0x002 (SYN)
```

```
Window size value: 16384
```

```
[Calculated window size: 16384]
```

```
Checksum: 0x2025 [unverified]
```

```
[Checksum Status: Unverified]
```

```
Urgent pointer: 0
```

```
Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
```

```
Maximum segment size: 1460 bytes
```

```
Kind: Maximum Segment Size (2)
```

```
Length: 4
```

MSS Value: 1460

Window scale: 0 (multiply by 1)

End of Option List (EOL)

TCP SYN、R5からACK:

! - TCP SYN, ACK sourced from R5

238 2696.702792 10.5.6.5 10.5.6.6 TCP 62 179 47007 [SYN, ACK] Seq=0
Ack=1 Win=16384 Len=0 **MSS=1460** WS=1

Frame 238: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:ad:51:8f (fa:16:3e:ad:51:8f), Dst: fa:16:3e:85:3d:c2
(fa:16:3e:85:3d:c2)

Internet Protocol Version 4, Src: 10.5.6.5, Dst: 10.5.6.6

Transmission Control Protocol, Src Port: 179, Dst Port: 47007, Seq: 0, Ack: 1, Len: 0

Source Port: 179

Destination Port: 47007

[Stream index: 10]

[TCP Segment Len: 0]

Sequence number: 0 (relative sequence number)

Acknowledgment number: 1 (relative ack number)

Header Length: 28 bytes

Flags: 0x012 (SYN, ACK)

Window size value: 16384

[Calculated window size: 16384]

Checksum: 0x7078 [unverified]

[Checksum Status: Unverified]

Urgent pointer: 0

Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)

Maximum segment size: 1460 bytes

Kind: Maximum Segment Size (2)

Length: 4

MSS Value: 1460

Window scale: 0 (multiply by 1)

End of Option List (EOL)

R6に表示されるTCPセッションの詳細 - ACTIVE:

! - as seen on R6 - Active

RP/0/0/CPU0:R6#show tcp detail pcb 0x1215761c

Fri Jan 8 10:15:20.351 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Fri Jan 8 10:10:04 2021

PCB 0x1215761c, SO 0x12162aac, TCPCB 0x12156f64, vrfid 0x60000000,

Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 103

Local host: 10.5.6.6, Local port: 47007 (Local App PID: 1011918)

Foreign host: 10.5.6.5, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	10	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	7	5	0
KeepAlive	1	0	0

```
PmtuAger          0          0          0
GiveUp            0          0          0
Throttle         0          0          0
```

```
iss: 3949093168  snduna: 3949093359  sndnxt: 3949093359
sndmax: 3949093359  sndwnd: 32631      sndcwnd: 4380
irs: 54439005    rcvnxt: 54439196   rcvwnd: 32650   rcvadp: 54471846
```

```
SRTT: 75 ms,  RTTO: 459 ms,  RTV: 384 ms,  KRTT: 0 ms
minRTT: 9 ms,  maxRTT: 239 ms
```

```
ACK hold time: 200 ms,  Keepalive time: 0 sec,  SYN waittime: 30 sec
Giveup time: 0 ms,  Retransmission retries: 0,  Retransmit forever: FALSE
Connect retries remaining: 30,  connect retry interval: 50 secs
```

```
State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale
```

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

```
Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none
Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer   : Low/High watermark 2048/24576, Notify threshold 0
```

```
PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40  PD ctx: size: 0  data:
Num Labels: 0  Label Stack:
```

RP/0/0/CPU0:R6#

R5に表示されるTCPセッションの詳細 : パッシブ :

! - as seen on R5 - Passive

```
RP/0/0/CPU0:R5#show interfaces gigabitEthernet 0/0/0/0
Fri Jan  8 10:10:39.110 UTC
GigabitEthernet0/0/0/0 is up, line protocol is up
Interface state transitions: 1
Hardware is GigabitEthernet, address is fa16.3ead.518f (bia fa16.3ead.518f)
Internet address is 10.5.6.5/30
MTU 4474 bytes, BW 1000000 Kbit (Max: 1000000 Kbit)
<snip>
```

```
RP/0/0/CPU0:R5#show tcp detail pcb 0x121550fc
Fri Jan  8 10:14:20.105 UTC
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan  8 10:10:05 2021
```

```
PCB 0x121550fc, SO 0x12154e18, TCPCB 0x12154304, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 103
Local host: 10.5.6.5, Local port: 179 (Local App PID: 1044686)
Foreign host: 10.5.6.6, Foreign port: 47007
```

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	7	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	7	2	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 54439005 snduna: 54439177 sndnxt: 54439177
sndmax: 54439177 sndwnd: 32669 sndcwnd: 4380
irs: 3949093168 rcvnxt: 3949093340 rcvwnd: 32650 rcvadv: 3949125990

SRTT: 117 ms, RTTO: 570 ms, RTV: 453 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 4420, max MSS 4420

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R5#

TCPオプションの使用 – XRアクティブ

このドキュメントで前述したように、TCPオプション([TCP MD5](#)、[TCP selective-ack](#)、[TCP timestamps](#)など)の使用は、MSS計算に含まれる追加バイトに導かれるため、MSS計算に影響します。

このセクションと次の目的は、TCPオプションが存在する場合にピアによって行われるMSS計算を説明することです。例としてTCP MD5認証オプションを使用します。図に示すように、図2.4の参照シナリオを参照してください。

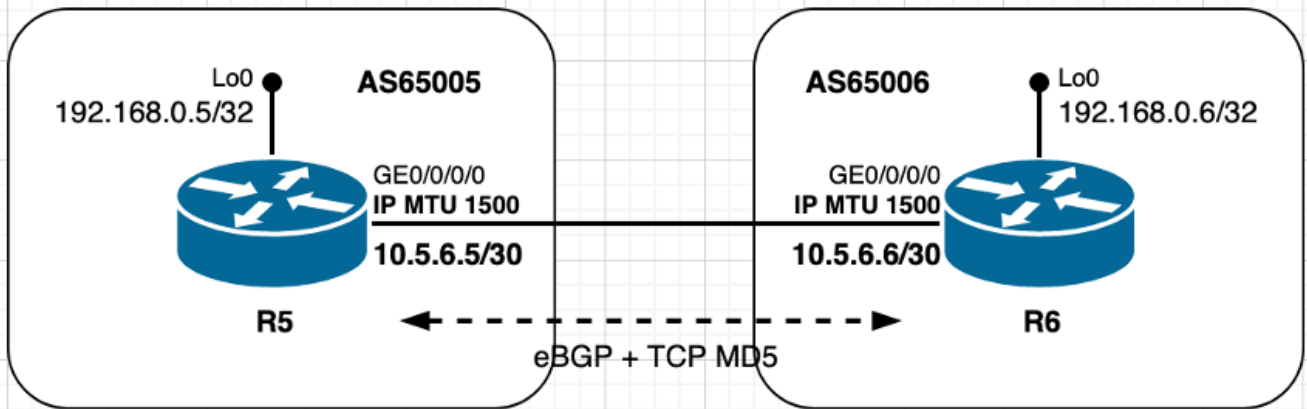


図2.4:TCPオプションの使用(MD5) - XRアクティブ。

このシナリオでは、両方のピアがデフォルトのIP MTU値を使用して直接接続され、ピアR6がTCPのアクティブな役割を果たします。すでにTCP MD5認証の設定と使用を共有しているため、オーバーヘッドが増加します。この特定のシナリオでのTCP MSSの計算は、次のように要約できます。

- 両方のノードが1500バイトのデフォルトIP MTUを使用します
- TCPパスMTUディスカバリーはデフォルトで無効になっています
- TCPピアが直接接続されている
- TCP MD5認証が両方のノードで有効になっている R6はBGP接続を管理しますR6は1436バイトのMSSでSYNを送信します 1500 (インターフェイスIP MTU) - 20(minTCP_H) - 20(minIP_H) - 24バイト (IOS XR TCPオプションオーバーヘッド) R5はSYNを送信し、MSSが1436バイトのACK [Received MSS;ローカル初期MSS]受信MSS 1436バイト。ローカル初期MSS 1460バイト最小のMSS値が両方のピアで使用されます

要約からわかるように、Cisco IOS XRの動作は、[RFC 879](#)および[RFC 6691](#)に厳密に準拠していません。[RFC 6691](#)では、TCPオプションはMSS計算で考慮されるべきではありません。

tcpヘッダー長に関する余分な要因のCisco IOS XRアカウントは、Cisco Bug ID [CSCvf20166](#):

"(..) XRがBGP接続を開始する際、BGPはまずソケットを作成し、次にMD5を含むソケットオプションを設定します。これによりtcp option header length = 24になり、初期MSSは1500 - 4 - 2 = 1436になります。これはピアに送信され、ピアはmin(1436, 1460) = 1436(..)を使用します。

R6から送信されたTCP SYN:

```
! - TCP SYN sourced from R6
```

```
430      5775.839420    10.5.6.6      10.5.6.5      TCP      82      24785  179 [SYN] Seq=0
Win=16384 Len=0 MSS=1436 WS=1
```

```
Frame 430: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
Ethernet II, Src: fa:16:3e:85:3d:c2 (fa:16:3e:85:3d:c2), Dst: fa:16:3e:ad:51:8f
(fa:16:3e:ad:51:8f)
```

```
Internet Protocol Version 4, Src: 10.5.6.6, Dst: 10.5.6.5
```

```
Transmission Control Protocol, Src Port: 24785, Dst Port: 179, Seq: 0, Len: 0
```

```
Source Port: 24785
```

```
Destination Port: 179
```

```
[Stream index: 14]
```

```

[TCP Segment Len: 0]
Sequence number: 0 (relative sequence number)
Acknowledgment number: 0
Header Length: 48 bytes
Flags: 0x002 (SYN)
Window size value: 16384
[Calculated window size: 16384]
Checksum: 0xd62b [unverified]
[Checksum Status: Unverified]
Urgent pointer: 0
Options: (28 bytes), Maximum segment size, Window scale, No-Operation (NOP), TCP MD5 signature, End of Option List (EOL)
    Maximum segment size: 1436 bytes
        Kind: Maximum Segment Size (2)
        Length: 4
        MSS Value: 1436
    Window scale: 0 (multiply by 1)
    No-Operation (NOP)
    TCP MD5 signature
    End of Option List (EOL)

```

TCP SYN、R5からACK:

! - TCP SYN, ACK sourced from R5

```

431      5775.845744      10.5.6.5      10.5.6.6      TCP      82      179  24785 [SYN, ACK] Seq=0
Ack=1 Win=16384 Len=0 MSS=1436 WS=1

```

```

Frame 431: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
Ethernet II, Src: fa:16:3e:ad:51:8f (fa:16:3e:ad:51:8f), Dst: fa:16:3e:85:3d:c2
(fa:16:3e:85:3d:c2)
Internet Protocol Version 4, Src: 10.5.6.5, Dst: 10.5.6.6
Transmission Control Protocol, Src Port: 179, Dst Port: 24785, Seq: 0, Ack: 1, Len: 0
    Source Port: 179
    Destination Port: 24785
    [Stream index: 14]
    [TCP Segment Len: 0]
    Sequence number: 0 (relative sequence number)
    Acknowledgment number: 1 (relative ack number)
    Header Length: 48 bytes
    Flags: 0x012 (SYN, ACK)
    Window size value: 16384
    [Calculated window size: 16384]
    Checksum: 0xe83d [unverified]
    [Checksum Status: Unverified]
    Urgent pointer: 0
    Options: (28 bytes), Maximum segment size, Window scale, No-Operation (NOP), TCP MD5 signature, End of Option List (EOL)
        Maximum segment size: 1436 bytes
            Kind: Maximum Segment Size (2)
            Length: 4
            MSS Value: 1436
        Window scale: 0 (multiply by 1)
        No-Operation (NOP)
        TCP MD5 signature
        End of Option List (EOL)

```

R6に表示されるTCPセッションの詳細 - ACTIVE:

! - as seen on R6 - Active

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1215761c
```


Fri Jan 8 11:14:13.599 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 11:01:21 2021

PCB 0x1215761c, SO 0x1216419c, TCPCB 0x121649fc, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 409
Local host: 10.5.6.6, Local port: 24785 (Local App PID: 1011918)
Foreign host: 10.5.6.5, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	17	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	14	13	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 1379482495 snduna: 1379482819 sndnxt: 1379482819
sndmax: 1379482819 sndwnd: 32498 sndcwnd: 4308
irs: 3750694052 rcvnxt: 3750694376 rcvwnd: 32517 rcvadv: 3750726893

SRTT: 55 ms, RTTO: 300 ms, RTV: 176 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 259 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 50 secs

State flags: none
Feature flags: MD5, Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1436, peer MSS 1436, min MSS 1436, max MSS 1436

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R6#
R5に表示されるTCPセッションの詳細 : パッシブ :

! - as seen on R5 - Passive

RP/0/0/CPU0:R5#show tcp detail pcb 0x12155d04

Fri Jan 8 11:12:51.984 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 11:01:22 2021

PCB 0x12155d04, SO 0x12154e18, TCPCB 0x12154304, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 409
Local host: 10.5.6.5, Local port: 179 (Local App PID: 1044686)
Foreign host: 10.5.6.6, Foreign port: 24785

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	14	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	14	3	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 3750694052 snduna: 3750694357 sndnxt: 3750694357
sndmax: 3750694357 sndwnd: 32536 sndcwnd: 4308
irs: 1379482495 rcvnx: 1379482800 rcvwnd: 32517 rcvadv: 1379515317
SRTT: 181 ms, RTTO: 443 ms, RTV: 262 ms, KRTT: 0 ms
minRTT: 29 ms, maxRTT: 219 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: MD5, Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1436, peer MSS 1436, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R5#

他のTCPオプションでも同様の動作が確認できます。このオプションを設定すると、オーバーヘッドが増加し、Cisco IOS XRでのMSSの計算に影響が及びます。TCPタイムスタンプとTCP選択的確認応答(SELL)オプションが設定されている場合のMSSの計算を文書化する同じシナリオと次

の例を検討してください。

R6(ACTIVE)で表示されるTCPセッションの詳細。TCPオプションのタイムスタンプと選択的ACKオプションが設定されています。

```
! - as seen on R6 - Active
! -- tcp timestamp configured
! -- 12 bytes of additional overhead
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1539c844
<snip>
Feature flags: Timestamp, Win Scale, Nagle
Request flags: Timestamp, Win Scale
```

```
Datagrams (in bytes): MSS 1448, peer MSS 1448, min MSS 1448, max MSS 1448
<snip>
```

```
! - as seen on R6 - Active
! -- tcp selective-ack configured
! -- 36 bytes of additional overhead
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1539df38
<snip>
Feature flags: Sack, Win Scale, Nagle
Request flags: Sack, Win Scale
```

```
Datagrams (in bytes): MSS 1424, peer MSS 1424, min MSS 1424, max MSS 1424
<snip>
```

```
! - as seen on R6 - Active
! -- tcp selective-ack and tcp timestamp configured
! -- 40 bytes of additional overhead
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1539e130
<snip>
State flags: none
Feature flags: Sack, Timestamp, Win Scale, Nagle
Request flags: Sack, Timestamp, Win Scale
```

```
Datagrams (in bytes): MSS 1420, peer MSS 1420, min MSS 1420, max MSS 1420
<snip>
```

```
! - as seen on R6 - Active
! -- MD5 and tcp selective-ack configured
! -- 36 bytes of additional overhead
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1539b3cc
<snip>
Feature flags: Sack, MD5, Win Scale, Nagle
Request flags: Sack, Win Scale
```

```
Datagrams (in bytes): MSS 1424, peer MSS 1424, min MSS 1424, max MSS 1424
<snip>
```

```
! - as seen on R6 - Active
! -- MD5 and tcp timestamp configured
! -- 36 bytes of additional overhead
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x15397b4c
<snip>
```

```
Feature flags: MD5, Timestamp, Win Scale, Nagle
Request flags: Timestamp, Win Scale
```

```
Datagrams (in bytes): MSS 1424, peer MSS 1424, min MSS 1424, max MSS 1424
<snip>
```

```
! - as seen on R6 - Active
! -- MD5, tcp timestamp, and tcp selective-ack configured
! -- 40 bytes of additional overhead
```

```
RP/0/0/CPU0:R6#show tcp detail pcb 0x1539a4cc
```

```
<snip>
```

```
State flags: none
Feature flags: MD5, Timestamp, Win Scale, Nagle
Request flags: Timestamp, Win Scale
```

```
Datagrams (in bytes): MSS 1420, peer MSS 1420, min MSS 1420, max MSS 1420
<snip>
```

TCPオプションの使用：XRパッシブ

以前のシナリオから、初期MSS計算に関して受動的な役割を果たす場合のCisco IOS XRノードの明確な動作に気づいたことがあります。ノードはtcpオプションヘッダー長を考慮しません。このシナリオは、Cisco Bug ID (登録ユーザ専用)でも説明されている次の動作を強調することを目的としています。

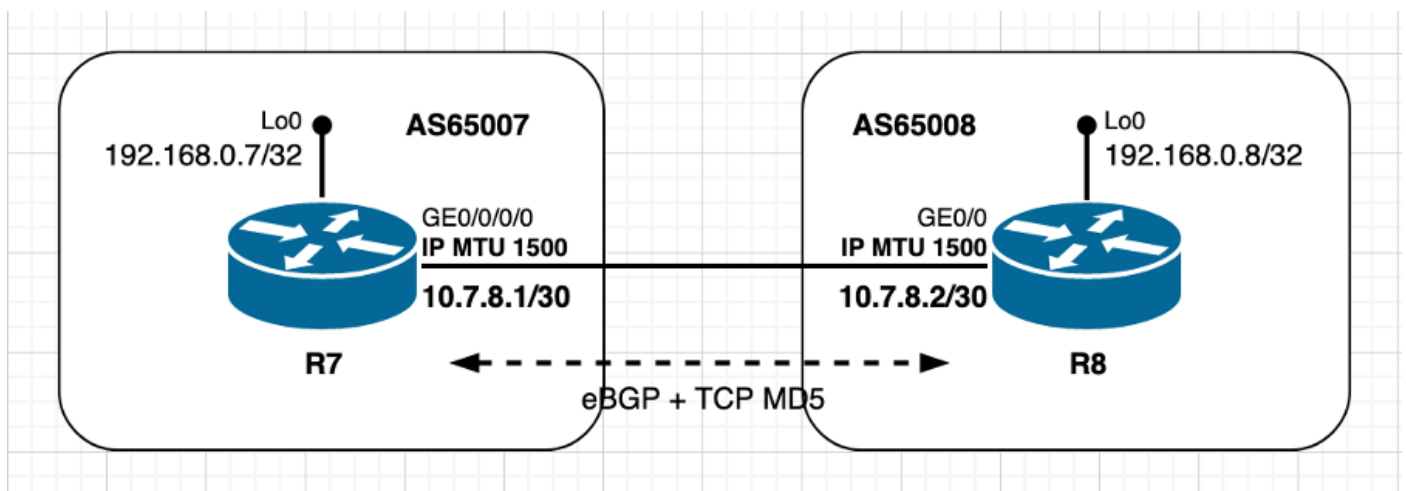
"(...) - ピアが接続を開始すると、初期MSSを1460として送信します。XR TCPはソケットやpcbなどを作成し、次の2つのアクションを順番に実行します。

- 最初に、tcpオプションヘッダー長を引いた後で初期MSSを計算します。これは'0'です。MD5オプションは、まだリスニングソケットからこのソケットに継承されていません。

- 次に、「MD5」とその他のオプションを継承し、「オプションヘッダーのバイト長」を24にします。

したがって、この場合、XR TCPは初期MSSを1460として送信するため、両方で使用されます。(...)"

このシナリオでは、アクティブなTCPピアR8はCisco IOSノードですが、この事実は、シナリオの要点に関する相違点や詳細を示すものではありません。ただし、前述のセクションのシナリオで示したように、Cisco IOS XRとは異なり、アクティブなTCPピアR8では、初期MSS計算にTCPオプションは考慮されないことに注意してください。



イメージ2.5:TCPオプションの使用(MD5) - XRパッシブ。

両方のピアはデフォルトのIP MTU値を使用し、直接接続されています。Cisco IOSピアR8はアクティブな役割を果たします。このシナリオでのTCP MSSの計算は、次のように要約できます。

- 両方のノードが1500バイトのデフォルトIP MTUを使用します
- Cisco IOS XR R7では、TCPパスMTUディスカバリはデフォルトで無効になっています
- Cisco IOS R8では、TCPパスMTUディスカバリがデフォルトで有効になっています
- TCPピアが直接接続されている
- TCP MD5認証が両方のノードで有効になっている IOS R8はBGP接続を管理しますIOS R8がSYNを1460バイトのMSSで送信 1500 (インターフェイスIP MTU) ~ 20(minTCP_H)~ 20(minIP_H)IOS XR R7はSYN、ACK、MSS 1460バイトを送信 [Received MSS;ローカル初期MSS]受信MSS 1460バイト。ローカル初期MSS 1460バイト最小のMSS値が両方のピアで使用されます

R8から送信されたTCP SYN - Cisco IOS:

```
! - TCP SYN sourced from R8
```

```
96      5.907127      10.7.8.2      10.7.8.1      TCP      78      52975  179 [SYN] Seq=0
Win=16384 Len=0  MSS=1460
```

```
Frame 96: 78 bytes on wire (624 bits), 78 bytes captured (624 bits) on interface 0
Ethernet II, Src: fa:16:3e:58:21:ba (fa:16:3e:58:21:ba), Dst: fa:16:3e:68:d9:e5
(fa:16:3e:68:d9:e5)
```

```
Internet Protocol Version 4, Src: 10.7.8.2, Dst: 10.7.8.1
```

```
Transmission Control Protocol, Src Port: 52975, Dst Port: 179, Seq: 0, Len: 0
```

```
Source Port: 52975
```

```
Destination Port: 179
```

```
[Stream index: 3]
```

```
[TCP Segment Len: 0]
```

```
Sequence number: 0 (relative sequence number)
```

```
Acknowledgment number: 0
```

```
Header Length: 44 bytes
```

```
Flags: 0x002 (SYN)
```

```
Window size value: 16384
```

```
[Calculated window size: 16384]
```

```
Checksum: 0xb612 [unverified]
```

```
[Checksum Status: Unverified]
```

```
Urgent pointer: 0
```

```
Options: (24 bytes), Maximum segment size, TCP MD5 signature, End of Option List (EOL)
```

```
Maximum segment size: 1460 bytes
```

```
Kind: Maximum Segment Size (2)
```

```
Length: 4
```

```
MSS Value: 1460
```

```
TCP MD5 signature
```

```
End of Option List (EOL)
```

TCP SYN、ACK送信元R7 - Cisco IOS XR:

```
! - TCP SYN,ACK sourced from R7
```

```
97      0.003446      10.7.8.1      10.7.8.2      TCP      78      179  52975 [SYN, ACK] Seq=0
Ack=1 Win=16384 Len=0  MSS=1460
```

```
Frame 97: 78 bytes on wire (624 bits), 78 bytes captured (624 bits) on interface 0
Ethernet II, Src: fa:16:3e:68:d9:e5 (fa:16:3e:68:d9:e5), Dst: fa:16:3e:58:21:ba
(fa:16:3e:58:21:ba)
```

```
Internet Protocol Version 4, Src: 10.7.8.1, Dst: 10.7.8.2
Transmission Control Protocol, Src Port: 179, Dst Port: 52975, Seq: 0, Ack: 1, Len: 0
  Source Port: 179
  Destination Port: 52975
  [Stream index: 3]
  [TCP Segment Len: 0]
  Sequence number: 0 (relative sequence number)
  Acknowledgment number: 1 (relative ack number)
  Header Length: 44 bytes
  Flags: 0x012 (SYN, ACK)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0xfb47 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (24 bytes), Maximum segment size, TCP MD5 signature, End of Option List (EOL)
    Maximum segment size: 1460 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1460
    TCP MD5 signature
    End of Option List (EOL)
```

R8に表示されるTCPセッションの詳細 : Cisco IOS - ACTIVE:

! - as seen from R8 - Cisco IOS

```
R8#show ip bgp neighbors
BGP neighbor is 10.7.8.1, remote AS 65007, external link
BGP version 4, remote router ID 192.168.0.7
BGP state = Established, up for 00:06:12
Last read 00:00:16, last write 00:00:16, hold time is 180, keepalive interval is 60 seconds
Neighbor sessions:
  1 active, is not multisession capable (disabled)
Neighbor capabilities:
  Route refresh: advertised and received(new)
  Four-octets ASN Capability: advertised and received
  Address family IPv4 Unicast: advertised and received
  Enhanced Refresh Capability: advertised
  Multisession Capability:
  Stateful switchover support enabled: NO for session 1
Message statistics:
  InQ depth is 0
  OutQ depth is 0

              Sent          Rcvd
Opens:                1            1
Notifications:       0            0
Updates:              1            1
Keepalives:           7            7
Route Refresh:        0            0
Total:                 9            9

Do log neighbor state changes (via global configuration)
Default minimum time between advertisement runs is 30 seconds

For address family: IPv4 Unicast
Session: 10.7.8.1
BGP table version 1, neighbor version 1/0
Output queue size : 0
Index 6, Advertise bit 0
6 update-group member
Slow-peer detection is disabled
Slow-peer split-update-group dynamic is disabled
```

	Sent	Rcvd
Prefix activity:	----	----
Prefixes Current:	0	0
Prefixes Total:	0	0
Implicit Withdraw:	0	0
Explicit Withdraw:	0	0
Used as bestpath:	n/a	0
Used as multipath:	n/a	0
Used as secondary:	n/a	0

	Outbound	Inbound
Local Policy Denied Prefixes:	-----	-----
Total:	0	0

Number of NLRI in the update sent: max 0, min 0

Last detected as dynamic slow peer: never

Dynamic slow peer recovered: never

Refresh Epoch: 1

Last Sent Refresh Start-of-rib: never

Last Sent Refresh End-of-rib: never

Last Received Refresh Start-of-rib: never

Last Received Refresh End-of-rib: never

	Sent	Rcvd
Refresh activity:	----	----
Refresh Start-of-RIB	0	0
Refresh End-of-RIB	0	0

Address tracking is enabled, the RIB does have a route to 10.7.8.1

Connections established 6; dropped 5

Last reset 00:06:18, due to BGP Notification received of session 1, Administrative Reset

External BGP neighbor configured for connected checks (single-hop no-disable-connected-check)

Interface associated: GigabitEthernet0/1 (peering address in same link)

Transport(tcp) path-mtu-discovery is enabled

Graceful-Restart is disabled

SSO is disabled

Connection state is ESTAB, I/O status: 1, unread input bytes: 0

Connection is ECN Disabled, Minimum incoming TTL 0, Outgoing TTL 1

Local host: 10.7.8.2, Local port: 52975

Foreign host: 10.7.8.1, Foreign port: 179

Connection tableid (VRF): 0

Maximum output segment queue size: 50

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

Event Timers (current time is 0x15DD97):

Timer	Starts	Wakeups	Next
Retrans	10	0	0x0
TimeWait	0	0	0x0
AckHold	9	5	0x0
SendWnd	0	0	0x0
KeepAlive	0	0	0x0
GiveUp	0	0	0x0
PmtuAger	1	0	0x195465
DeadWait	0	0	0x0
Linger	0	0	0x0
ProcessQ	0	0	0x0

iss: 1154289541 snduna: 1154289755 sndnxt: 1154289755

irs: 2149897425 rcvnxt: 2149897635

sndwnd: 32612 scale: 0 maxrcvwnd: 16384

rcvwnd: 16175 scale: 0 delrcvwnd: 209

SRTT: 737 ms, RTTO: 2506 ms, RTV: 1769 ms, KRTT: 0 ms

minRTT: 7 ms, maxRTT: 1000 ms, ACK hold: 200 ms

uptime: 372981 ms, Sent idletime: 16648 ms, Receive idletime: 16431 ms
Status Flags: active open
Option Flags: nagle, path mtu capable, **md5**
IP Precedence value : 6

Datagrams (max data segment is 1460 bytes):

Rcvd: 18 (out of order: 0), with data: 8, total data bytes: 209
Sent: 16 (retransmit: 0, fastretransmit: 0, partialack: 0, Second Congestion: 0), with data: 9,
total data bytes: 213

Packets received in fast path: 0, fast processed: 0, slow path: 0
fast lock acquisition failures: 0, slow path: 0

TCP Semaphore 0x0FBFA8A4 FREE

R8#

R7 - Cisco IOS XR - PASSIVEに表示されるTCPセッションの詳細 :

! - as seen from R7 - Cisco IOS XR

RP/0/0/CPU0:R7#show tcp detail pcb 0x12152e48

Wed Jan 13 13:03:43.363 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Wed Jan 13 12:58:16 2021

PCB 0x12152e48, SO 0x1213c130, TCPCB 0x12156060, vrfid 0x60000000,

Pak Prio: Medium, TOS: 192, TTL: 1, Hash index: 947

Local host: 10.7.8.1, Local port: 179 (Local App PID: 983244)

Foreign host: 10.7.8.2, Foreign port: 52975

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	8	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	8	7	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 2149897425 snduna: 2149897616 sndnxt: 2149897616
sndmax: 2149897616 sndwnd: 16194 sndcwnd: 4380
irs: 1154289541 rcvnxt: 1154289736 rcvwnd: 32631 rcvadv: 1154322367

SRTT: 125 ms, RTTO: 552 ms, RTV: 427 ms, KRTT: 0 ms

minRTT: 19 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec

Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE

Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none

Feature flags: **MD5**, Nagle

Request flags: none

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0


```

Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R7#

```

直接接続されていないTCPピア

ピアが直接接続されていない場合、TCP MSSの初期計算の方法は、このドキュメントの最初のセクションで説明したように変更されます。デフォルトのIP MTU値が設定されたすべてのピアとのiBGPセッションのシナリオは、MSS計算のウォークスルーに使用されます。

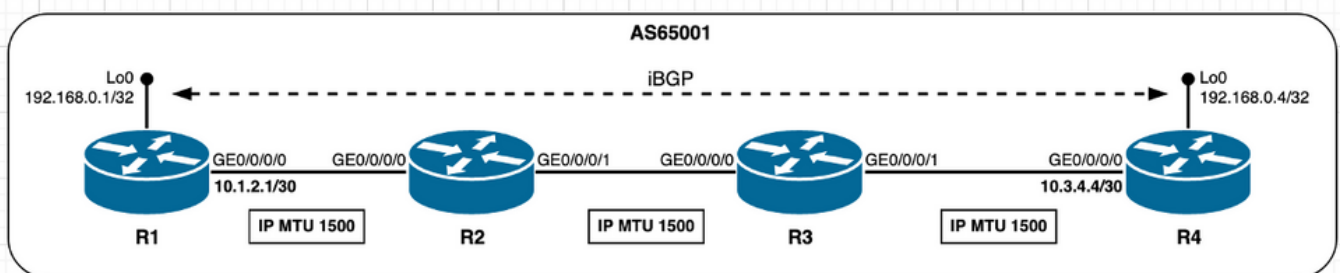


図2.6 : 直接接続されていないTCPピア – iBGP。

注意すべき重要な点は、TCPパスMTUディスカバリが無効で、ピアが直接接続されていない場合、Cisco IOS XRでは固定IP MTU値1280バイトが使用されることです。

前の図では、R4がアクティブな役割を果たし、TCP接続を管理しています。R4は、宛先ポート179でR1とのTCPセッションを開きます。両方のノードは、インターフェイスでデフォルトのIP MTU値を使用します。このシナリオのMSS計算は、次のように要約できます。

- すべてのノードが1500バイトのデフォルトIP MTUを使用します
- TCPパスMTUディスカバリはデフォルトで無効になっています
- TCPピアが直接接続されていない R4はBGP接続を管理しますR4は1240バイトのMSSでSYNを送信しますピアが直接接続されておらず、TCPパスMTUディスカバリが無効になっている場合、インターフェイスMTUは考慮されませんCisco IOS XRの設計では、1280バイトがTCP_DEFAULT_MTUと見なされます1280(TCP_DEFAULT_MTU) ~ 20(minTCP_H) ~ 20(minIP_H)R1は1240バイトのMSSでSYN、ACKを送信します [Received MSS;ローカル初期MSS]受信MSS 1240バイト。ローカル初期MSS 1240バイト最小のMSS値が両方のピアで使用されます

R4から送信されたTCP SYN:

```
! - TCP SYN sourced from R4
```

194 434.274181 192.168.0.4 192.168.0.1 TCP 62 37740 179 [SYN] Seq=0 Win=16384
Len=0 **MSS=1240** WS=1

Frame 194: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:d7:7e:f6 (fa:16:3e:d7:7e:f6), Dst: fa:16:3e:8f:8f:54
(fa:16:3e:8f:8f:54)

Internet Protocol Version 4, Src: 192.168.0.4, Dst: 192.168.0.1

Transmission Control Protocol, Src Port: 37740, Dst Port: 179, Seq: 0, Len: 0

Source Port: 37740

Destination Port: 179

[Stream index: 7]

[TCP Segment Len: 0]

Sequence number: 0 (relative sequence number)

Acknowledgment number: 0

Header Length: 28 bytes

Flags: 0x002 (SYN)

Window size value: 16384

[Calculated window size: 16384]

Checksum: 0x8643 [unverified]

[Checksum Status: Unverified]

Urgent pointer: 0

Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)

Maximum segment size: 1240 bytes

Kind: Maximum Segment Size (2)

Length: 4

MSS Value: 1240

Window scale: 0 (multiply by 1)

End of Option List (EOL)

TCP SYN、R1からACK:

! - TCP SYN,ACK sourced from R1

195 434.277985 192.168.0.1 192.168.0.4 TCP 62 179 37740 [SYN, ACK] Seq=0 Ack=1
Win=16384 Len=0 **MSS=1240** WS=1

Frame 195: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:8f:8f:54 (fa:16:3e:8f:8f:54), Dst: fa:16:3e:d7:7e:f6
(fa:16:3e:d7:7e:f6)

Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4

Transmission Control Protocol, Src Port: 179, Dst Port: 37740, Seq: 0, Ack: 1, Len: 0

Source Port: 179

Destination Port: 37740

[Stream index: 7]

[TCP Segment Len: 0]

Sequence number: 0 (relative sequence number)

Acknowledgment number: 1 (relative ack number)

Header Length: 28 bytes

Flags: 0x012 (SYN, ACK)

Window size value: 16384

[Calculated window size: 16384]

Checksum: 0xd8f7 [unverified]

[Checksum Status: Unverified]

Urgent pointer: 0

Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)

Maximum segment size: 1240 bytes

Kind: Maximum Segment Size (2)

Length: 4

MSS Value: 1240

Window scale: 0 (multiply by 1)

End of Option List (EOL)

R4に表示されるTCPセッションの詳細 - ACTIVE:

! - as seen on R4 - Active

RP/0/0/CPU0:R4#show tcp detail pcb 0x12154d3c

Fri Jan 8 12:32:41.096 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Fri Jan 8 12:17:46 2021

PCB 0x12154d3c, SO 0x12154460, TCPCB 0x1215486c, vrfid 0x60000000,

Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 1577

Local host: 192.168.0.4, Local port: 37740 (Local App PID: 1052958)

Foreign host: 192.168.0.1, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	19	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	16	15	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 2075436506 snduna: 2075436868 sndnxt: 2075436868
sndmax: 2075436868 sndwnd: 32460 sndcwnd: 3720
irs: 4238127261 rcvnxt: 4238127623 rcvwnd: 32479 rcvadv: 4238160102

SRTT: 65 ms, RTTO: 300 ms, RTV: 40 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 30 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R4#

R1に表示されるTCPセッションの詳細：パッシブ：

! - as seen on R1 - Passive

RP/0/0/CPU0:R1#show tcp detail pcb 0x12155390

Fri Jan 8 12:23:52.041 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Fri Jan 8 12:17:43 2021

PCB 0x12155390, SO 0x121573e4, TCPCB 0x12156948, vrfid 0x60000000,

Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 1577

Local host: 192.168.0.1, Local port: 179 (Local App PID: 983326)

Foreign host: 192.168.0.4, Foreign port: 37740

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	9	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	9	1	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 4238127261 snduna: 4238127471 sndnxt: 4238127471

sndmax: 4238127471 sndwnd: 32631 sndcwnd: 3720

irs: 2075436506 rcvnxt: 2075436716 rcvwnd: 32612 rcvadv: 2075469328

SRTT: 144 ms, RTTO: 578 ms, RTV: 434 ms, KRTT: 0 ms

minRTT: 19 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec

Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE

Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none

Feature flags: Win Scale, Nagle

Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0

Timestamp option: recent 0, recent age 0, last ACK sent 0

Sack blocks {start, end}: none

Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO

Socket states: SS_ISCONNECTED, SS_PRIV

Socket receive buffer states: SB_DEL_WAKEUP

Socket send buffer states: SB_DEL_WAKEUP

Socket receive buffer: Low/High watermark 1/32768

Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:

#PDU's in buffer: 0

FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:

Num Labels: 0 Label Stack:

直接接続されていないTCPピア – TCPオプションの使用(MD5)

直接接続されていないピアシナリオやTCP MD5認証を使用している場合、前述のテストケースやシナリオと基本的な違いはありません。TCP MD5認証で前述したように、Cisco IOS XRは追加のオーバーヘッドを考慮し、初期MSS値も同様に反映します。TCP MSSの計算に対するTCPオプションの影響の詳細については、前のセクション「TCPオプションの使用：XRアクティブおよびTCPオプションの使用：XRパッシブ」を参照してください。

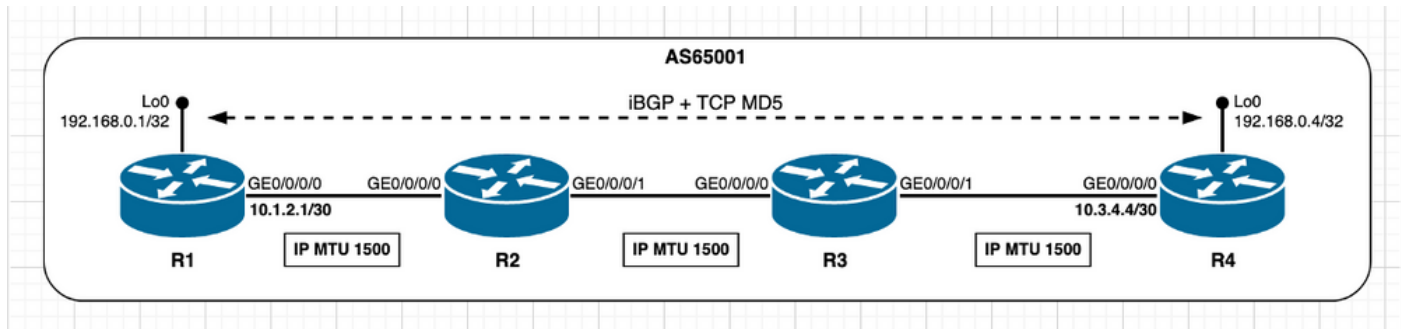


図2.7：直接接続されていないTCPピア – iBGP + TCP MD5。

このシナリオでのTCP MSSの計算は、次のように要約できます。

- すべてのノードが1500バイトのデフォルトIP MTUを使用します
- TCPパスMTUディスカバリはデフォルトで無効になっています
- TCPピアが直接接続されていない R4はBGP接続を管理します宛先R1が直接接続されていないR4は1216バイトのMSSでSYNを送信しますピアが直接接続されておらず、TCPパスMTUディスカバリが無効になっている場合、インターフェイスMTUは考慮されません設計上1280バイトはTCP_DEFAULT_MTUと見なされます1280(TCP_DEFAULT_MTU) - 20(minTCP_H) - 20(minIP_H) - 24バイト(IOS XR TCP Options Overhead)R1は1216バイトのMSSでSYN、ACKを送信します [Received MSS;ローカル初期MSS]受信MSS 1216バイト。ローカル初期MSS 1240バイト最小のMSS値が両方のピアで使用されます

R4から送信されたTCP SYN:

! - TCP SYN sourced from R4

```
3425  3.691042      192.168.0.4 192.168.0.1 TCP      82      42135  179 [SYN] Seq=0 Win=16384
Len=0 MSS=1216 WS=1
```

```
Frame 3425: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
Ethernet II, Src: fa:16:3e:d7:7e:f6 (fa:16:3e:d7:7e:f6), Dst: fa:16:3e:8f:8f:54
(fa:16:3e:8f:8f:54)
```

```
Internet Protocol Version 4, Src: 192.168.0.4, Dst: 192.168.0.1
```

```
Transmission Control Protocol, Src Port: 42135, Dst Port: 179, Seq: 0, Len: 0
```

```
Source Port: 42135
```

```
Destination Port: 179
```

```
[Stream index: 10]
```

```
[TCP Segment Len: 0]
```

```
Sequence number: 0 (relative sequence number)
```

```
Acknowledgment number: 0
```

```
Header Length: 48 bytes
```

```
Flags: 0x002 (SYN)
```

```
Window size value: 16384
```

```
[Calculated window size: 16384]
Checksum: 0xc503 [unverified]
[Checksum Status: Unverified]
Urgent pointer: 0
Options: (28 bytes), Maximum segment size, Window scale, No-Operation (NOP), TCP MD5
signature, End of Option List (EOL)
  Maximum segment size: 1216 bytes
    Kind: Maximum Segment Size (2)
    Length: 4
    MSS Value: 1216
  Window scale: 0 (multiply by 1)
  No-Operation (NOP)
  TCP MD5 signature
  End of Option List (EOL)
```

TCP SYN、R1からACK:

! - TCP SYN,ACK sourced from R1

```
3426  0.004186      192.168.0.1 192.168.0.4 TCP      82      179  42135 [SYN, ACK] Seq=0 Ack=1
Win=16384 Len=0 MSS=1216 WS=1
```

```
Frame 3426: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
Ethernet II, Src: fa:16:3e:8f:8f:54 (fa:16:3e:8f:8f:54), Dst: fa:16:3e:d7:7e:f6
(fa:16:3e:d7:7e:f6)
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
Transmission Control Protocol, Src Port: 179, Dst Port: 42135, Seq: 0, Ack: 1, Len: 0
  Source Port: 179
  Destination Port: 42135
  [Stream index: 10]
  [TCP Segment Len: 0]
  Sequence number: 0      (relative sequence number)
  Acknowledgment number: 1      (relative ack number)
  Header Length: 48 bytes
  Flags: 0x012 (SYN, ACK)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0xbb05 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (28 bytes), Maximum segment size, Window scale, No-Operation (NOP), TCP MD5
signature, End of Option List (EOL)
    Maximum segment size: 1216 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1216
    Window scale: 0 (multiply by 1)
    No-Operation (NOP)
    TCP MD5 signature
    End of Option List (EOL)
```

R4に表示されるTCPセッションの詳細 - ACTIVE:

! - as seen from R4 - Active

```
RP/0/0/CPU0:R4#show tcp detail pcb 0x12154490
Tue Jan 12 14:37:32.097 UTC
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Tue Jan 12 14:27:42 2021

PCB 0x12154490, SO 0x12155014, TCPCB 0x12155a84, vrfid 0x60000000,
```

Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 1876
Local host: 192.168.0.4, Local port: 42135 (Local App PID: 1052958)
Foreign host: 192.168.0.1, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	14	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	11	9	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 3124761989 snduna: 3124763317 sndnxt: 3124763317
sndmax: 3124763317 sndwnd: 32711 sndcwnd: 3648
irs: 1090344992 rcvnxt: 1090346320 rcvwnd: 32730 rcvadv: 1090379050

SRTT: 28 ms, RTTO: 300 ms, RTV: 57 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 30 secs

State flags: none
Feature flags: MD5, Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1216, peer MSS 1216, min MSS 1216, max MSS 1216

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R4#

R1に表示されるTCPセッションの詳細: パッシブ:

! - as seen from R1 - Passive

RP/0/0/CPU0:R1#show tcp detail pcb 0x12168df4
Tue Jan 12 14:36:38.860 UTC
=====
Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Tue Jan 12 14:27:32 2021

PCB 0x12168df4, SO 0x12156bf8, TCPCB 0x12157a44, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 1876
Local host: 192.168.0.1, Local port: 179 (Local App PID: 983326)
Foreign host: 192.168.0.4, Foreign port: 42135

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	12	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	12	1	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 1090344992 snduna: 1090346320 sndnxt: 1090346320
sndmax: 1090346320 sndwnd: 32730 sndcwnd: 3648
irs: 3124761989 rcvnxt: 3124763317 rcvwnd: 32711 rcvadv: 3124796028

SRTT: 150 ms, RTTO: 558 ms, RTV: 408 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: MD5, Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1216, peer MSS 1216, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

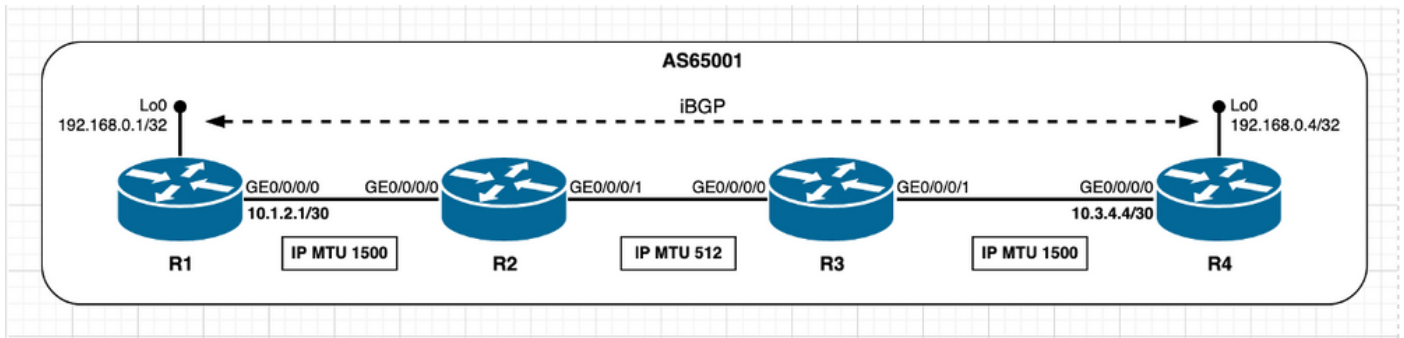
Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R1#

直接接続されていないTCPピア : パスセグメントのIP MTUが小さい

次のシナリオでは、目標は、IP MTUが低い中間パスセグメントがデフォルトの状態にある場合に何が起きるかを確認し、結論を出すことです。これは、TCP PMTUDが無効であることを意味します。この画像を参照してください。



イメージ2.8:R2/R3パスセグメントのIP MTUが小さい。

最初のシナリオでは、BGP情報が最小限であると考えます。つまり、BGPピア間で交換する必要があるものは、最小パスMTU 512バイトに収まるIPパケットで実現できます。この前提に基づき、MSS計算は、「直接接続されていないTCPピア」セクションで説明されているように行われま
す。R1とR4の両方が1240バイトのMSS値を選択します。

R4に表示されるTCPセッションの詳細 – ACTIVE:

! - as seen from R4 - Active

```
RP/0/0/CPU0:R4#show tcp detail pcb 0x15390fe8
```

```
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Wed May 12 12:09:48 2021
```

```
PCB 0x15390fe8, SO 0x15391a7c, TCPCB 0x15391368, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 835
Local host: 192.168.0.4, Local port: 39046 (Local App PID: 1196319)
Foreign host: 192.168.0.1, Foreign port: 179
(Local App PID/instance/SPL_APP_ID: 1196319/1/0)
```

```
Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)
```

Timer	Starts	Wakeups	Next(msec)
Retrans	1267	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	1280	1235	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

```
iss: 1991226354 snduna: 1991250450 sndnxt: 1991250450
sndmax: 1991250450 sndwnd: 32578 sndcwnd: 2480
irs: 4276699304 rcvnxt: 4276746737 rcvwnd: 31568 rcvadp: 4276778305
```

```
SRTT: 213 ms, RTTO: 300 ms, RTV: 54 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 269 ms
```

```
ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 10, connect retry interval: 30 secs
```

```
State flags: none
Feature flags: Win Scale, Nagle
```

Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

<snip>

R1に表示されるTCPセッションの詳細：パッシブ：

! - as seen from R1 - Passive

RP/0/0/CPU0:R1#show tcp detail pcb 0x15393770

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Wed May 12 12:09:46 2021

PCB 0x15393770, SO 0x15392224, TCPCB 0x153928cc, vrfid 0x60000000,

Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 835

Local host: 192.168.0.1, Local port: 179 (Local App PID: 1192224)

Foreign host: 192.168.0.4, Foreign port: 39046

(Local App PID/instance/SPL_APP_ID: 1192224/1/0)

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	1280	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	1264	1213	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 4276699304 snduna: 4276746718 sndnxt: 4276746718
sndmax: 4276746718 sndwnd: 31587 sndcwnd: 3720
irs: 1991226354 rcvnxt: 1991250431 rcvwnd: 32597 rcvadv: 1991283028

SRTT: 202 ms, RTTO: 355 ms, RTV: 153 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 309 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

<snip>

BGPセッションが確立された状態で、最小パスMTUが512バイトよりも大きいサイズのBGP Updateメッセージがトリガーされることを検討してください。出力からわかるように、Cisco IOS XRはBGPアップデートメッセージを使用してdf-bitを設定しません。これは、中間ノードでのパケット断片化を犠牲にしてBGP情報が送信されることを意味します。

R1が送信元とするBGPアップデート：パッシブ：

! - as seen from R1 - Passive - BGP UPDATE

! - Note Total Length of 1097 bytes higher than the IP MTU value of 512 bytes at R2-R3 path

segment

23 3.450878 192.168.0.1 192.168.0.4 BGP 1111 UPDATE Message

Frame 23: 1111 bytes on wire (8888 bits), 1111 bytes captured (8888 bits) on interface 0
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
(fa:16:3e:5c:f1:80)

Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4

0100 = Version: 4

.... 0101 = Header Length: 20 bytes (5)

Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)

Total Length: 1097

Identification: 0x5841 (22593)

Flags: 0x00

0... = Reserved bit: Not set

.0.. = Don't fragment: Not set

..0. = More fragments: Not set

Fragment offset: 0

Time to live: 255

Protocol: TCP (6)

Header checksum: 0x54a4 [validation disabled]

[Header checksum status: Unverified]

Source: 192.168.0.1

Destination: 192.168.0.4

[Source GeoIP: Unknown]

[Destination GeoIP: Unknown]

Transmission Control Protocol, Src Port: 179, Dst Port: 39046, Seq: 20, Ack: 20, Len: 1057

Border Gateway Protocol - UPDATE Message

Marker: ffffffffffffffffffffffffffffffffff

Length: 1057

Type: UPDATE Message (2)

Withdrawn Routes Length: 0

Total Path Attribute Length: 1034

Path attributes

Path Attribute - MP_REACH_NLRI

Path Attribute - ORIGIN: INCOMPLETE

Path Attribute - AS_PATH: empty

Path Attribute - MULTI_EXIT_DISC: 0

Path Attribute - LOCAL_PREF: 100

ノードR1から送信されたBGP Updateメッセージのフラグメンテーションは、ノードR2で発生します。これは、R2インターフェイスGE0/0/0/1で行われたトラフィックキャプチャで確認できません。

ノードR2でのIPフラグメンテーション :

! - as seen from R2 - GE0/0/0/1

! - Node R2 fragments original packet in three distinct packets

4 1.334852 192.168.0.1 192.168.0.4 BGP 522 UPDATE Message

5 0.000289 192.168.0.1 192.168.0.4 IPv4 522 Fragmented IP protocol (proto=TCP 6,
off=488, ID=7b41)

6 0.000122 192.168.0.1 192.168.0.4 IPv4 135 Fragmented IP protocol (proto=TCP 6,
off=976, ID=7b41)

! - Captured frame details

Frame 4: 522 bytes on wire (4176 bits), 522 bytes captured (4176 bits) on interface 0
Ethernet II, Src: fa:16:3e:61:25:f0 (fa:16:3e:61:25:f0), Dst: fa:16:3e:23:ab:27
(fa:16:3e:23:ab:27)

Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4

0100 = Version: 4

.... 0101 = Header Length: 20 bytes (5)
Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
Total Length: 508
Identification: 0x7b41 (31553)
Flags: 0x01 (More Fragments)
 0... = Reserved bit: Not set
 .0.. = Don't fragment: Not set
 ..1. = **More fragments: Set**
Fragment offset: 0
Time to live: 254
Protocol: TCP (6)
Header checksum: 0x14f1 [validation disabled]
[Header checksum status: Unverified]
Source: 192.168.0.1
Destination: 192.168.0.4
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Transmission Control Protocol, Src Port: 179, Dst Port: 39046, Seq: 4276759681, Ack: 1991250830
Border Gateway Protocol - UPDATE Message
<snip>

Frame 5: 522 bytes on wire (4176 bits), 522 bytes captured (4176 bits) on interface 0
Ethernet II, Src: fa:16:3e:61:25:f0 (fa:16:3e:61:25:f0), Dst: fa:16:3e:23:ab:27
(fa:16:3e:23:ab:27)

Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
 0100 = Version: 4
 0101 = Header Length: 20 bytes (5)
 Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
 Total Length: 508
 Identification: 0x7b41 (31553)
 Flags: 0x01 (More Fragments)
 0... = Reserved bit: Not set
 .0.. = Don't fragment: Not set
 ..1. = **More fragments: Set**
 Fragment offset: 488
 Time to live: 254
 Protocol: TCP (6)
 Header checksum: 0x14b4 [validation disabled]
 [Header checksum status: Unverified]
 Source: 192.168.0.1
 Destination: 192.168.0.4
 [Source GeoIP: Unknown]
 [Destination GeoIP: Unknown]
Data (488 bytes)
<snip>

Frame 6: 135 bytes on wire (1080 bits), 135 bytes captured (1080 bits) on interface 0
Ethernet II, Src: fa:16:3e:61:25:f0 (fa:16:3e:61:25:f0), Dst: fa:16:3e:23:ab:27
(fa:16:3e:23:ab:27)

Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
 0100 = Version: 4
 0101 = Header Length: 20 bytes (5)
 Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
 Total Length: 121
 Identification: 0x7b41 (31553)
 Flags: 0x00
 0... = Reserved bit: Not set
 .0.. = Don't fragment: Not set
 ..0. = **More fragments: Not set**
 Fragment offset: 976
 Time to live: 254
 Protocol: TCP (6)
 Header checksum: 0x35fa [validation disabled]
 [Header checksum status: Unverified]

```

Source: 192.168.0.1
Destination: 192.168.0.4
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Data (101 bytes)
<snip>

```

シナリオ : TCP PMTUD有効

PMTUDの有効化

PMTUDが有効になると、ピアが直接接続されているかどうかにかかわらず、MSSの初期計算では常に出カインターフェイスのIP MTUが考慮されます。

このシナリオでは、PMTUDが有効な場合に予想される動作について考察します。ここでは、Cisco IOS XRノードR4がアクティブな役割を果たし、TCP接続を管理し、Cisco IOS XRノードR1とのTCPセッションを宛先ポート179で開きます。どちらのノードもインターフェイスのデフォルトIP MTU値を使用します。

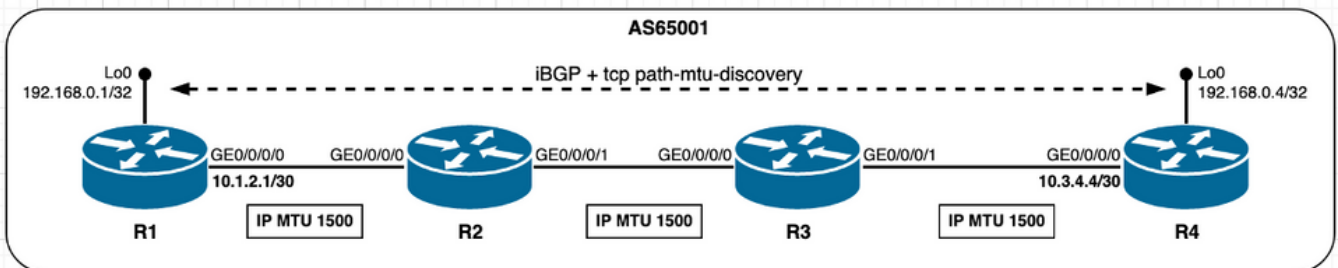


図3.1 - TCP PMTUD有効

このシナリオのMSS計算は、次のように要約できます。

- すべてのノードが1500バイトのデフォルトIP MTUを使用します
- TCPパスMTUディスカバリが有効になっている
- TCPピアが直接接続されていない R4はBGP接続を管理しますR4は1460バイトのMSSでSYNを送信します 1500 (インターフェイスIP MTU) ~ 20(minTCP_H) ~ 20(minIP_H)R1は1460バイトのMSSでSYN、ACKを送信します [Received MSS;ローカル初期MSS]受信MSS 1460バイト。ローカル初期MSS 1460バイト最小のMSS値が両方のピアで使用されます

enable PMTUDから導入された動作変更を強調表示するために、次の出力はイベントのシーケンスを示しています。

1. PMTUDが無効になっているデフォルトシナリオでの確立されたTCPセッションの初期状態
2. PMTUDが設定され、TCPピアR4とR1の両方で有効になっています。
3. TCPセッションが再起動し、MSS計算が行われ、TCP PMTUDの影響を受けます。

R4 - ACTIVE - TCP PMTUD disabled (デフォルト) に表示されます。

```

! - as seen on R4 - Active
! - TCP path mtu discovery disabled (default)
! - TCP session initial state

```

RP/0/0/CPU0:R4#show tcp detail pcb 0x121536c8

Fri Jan 8 16:06:30.237 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Fri Jan 8 16:05:15 2021

PCB 0x121536c8, SO 0x12155370, TCPCB 0x12154f64, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 376
Local host: 192.168.0.4, Local port: 20155 (Local App PID: 1052958)
Foreign host: 192.168.0.1, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	6	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	3	2	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 357400981 snduna: 357401257 sndnxt: 357401257
sndmax: 357401257 sndwnd: 32546 sndcwnd: 3720
irs: 524019443 rcvnxt: 524019719 rcvwnd: 32565 rcvadv: 524052284

SRTT: 72 ms, RTTO: 416 ms, RTV: 344 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 30 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R4#

R1 - PASSIVE - TCP PMTUD disabled (デフォルト) に表示されます。

! - as seen on R1 - Passive
! - TCP path mtu discovery disabled (default)
! - TCP session initial state

RP/0/0/CPU0:R1#show tcp detail pcb 0x12157020

Fri Jan 8 16:05:52.868 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 16:05:12 2021

PCB 0x12157020, SO 0x121565ac, TCPCB 0x121560ec, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 376
Local host: 192.168.0.1, Local port: 179 (Local App PID: 983326)
Foreign host: 192.168.0.4, Foreign port: 20155

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	3	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	3	1	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 524019443 snduna: 524019700 sndnxt: 524019700
sndmax: 524019700 sndwnd: 32584 sndcwnd: 3720
irs: 357400981 rcvnxt: 357401238 rcvwnd: 32565 rcvadv: 357433803

SRTT: 46 ms, RTTO: 300 ms, RTV: 249 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R1#

R4 - ACTIVE - TCP PMTUDが有効に設定されている場合：

! - 'debug tcp pmtud' output on R4
! - tcp path mtu discovery enabled and uses default Path MTU aging timer (10 min / 600000 msec)

RP/0/0/CPU0:Jan 8 16:09:28.285 : tcp[399]: [t21] Try to enable path MTU discovery(neww age timer: 10 min)

RP/0/0/CPU0:Jan 8 16:09:28.285 : tcp[399]: [t21] Path mtu is ON (age-timer: 10)

! - as seen on R4 - Active

! - TCP PMTUD is enabled

RP/0/0/CPU0:R4#show tcp detail pcb 0x121536c8

Fri Jan 8 16:11:00.138 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Fri Jan 8 16:05:15 2021

PCB 0x121536c8, SO 0x12155370, TCPCB 0x12154f64, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 376
Local host: 192.168.0.4, Local port: 20155 (Local App PID: 1052958)
Foreign host: 192.168.0.1, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	10	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	7	4	0
KeepAlive	1	0	0
PmtuAger	1	0	508096
GiveUp	0	0	0
Throttle	0	0	0

iss: 357400981 snduna: 357401333 sndnxt: 357401333
sndmax: 357401333 sndwnd: 32470 sndcwnd: 3720
irs: 524019443 rcvnx: 524019795 rcvwnd: 32489 rcvadv: 524052284

SRTT: 116 ms, RTTO: 578 ms, RTV: 462 ms, KRTT: 0 ms

minRTT: 9 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec

Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE

Connect retries remaining: 30, connect retry interval: 30 secs

State flags: PMTU ager

Feature flags: Win Scale, Nagle, **Path MTU**

Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0

Timestamp option: recent 0, recent age 0, last ACK sent 0

Sack blocks {start, end}: none

Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO

Socket states: SS_ISCONNECTED, SS_PRIV

Socket receive buffer states: SB_DEL_WAKEUP

Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:

#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R4#

R1 - PASSIVE - TCP PMTUDが有効に設定されている場合 :

! - 'debug tcp pmtud' output on R1
! - tcp path mtu discovery is enabled and uses default Path MTU aging timer (10 min / 60000 msec)

RP/0/0/CPU0:Jan 8 16:09:25.214 : tcp[399]: [t21] Try to enable path MTU discovery(neww age timer: 10 min)

RP/0/0/CPU0:Jan 8 16:09:25.214 : tcp[399]: [t21] Path mtu is ON (age-timer: 10)

! - as seen on R1 - Passive
! - TCP PMTUD is enabled

RP/0/0/CPU0:R1#show tcp detail pcb 0x12157020

Fri Jan 8 16:10:03.101 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 16:05:12 2021

PCB 0x12157020, SO 0x121565ac, TCPCB 0x121560ec, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 376
Local host: 192.168.0.1, Local port: 179 (Local App PID: 983326)
Foreign host: 192.168.0.4, Foreign port: 20155

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	7	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	7	4	0
KeepAlive	1	0	0
PmtuAger	1	0	562042
GiveUp	0	0	0
Throttle	0	0	0

iss: 524019443 snduna: 524019776 sndnxt: 524019776
sndmax: 524019776 sndwnd: 32508 sndcwnd: 3720
irs: 357400981 rcvnxt: 357401314 rcvwnd: 32489 rcvadv: 357433803

SRTT: 95 ms, RTTO: 528 ms, RTV: 433 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: PMTU ager
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 1240, peer MSS 1240, min MSS 1240, max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R1#

PMTU agerタイマーの動作に注意してください。

! - Note PmtuAger timer initial value is 10min
! - but after initial interval expires then it expires every 2min
! - As seen from 'debug tcp pmtud' output
! - TCP PMTUD is enabled

RP/0/0/CPU0:Jan 8 16:09:25.214 : tcp[399]: [t21] Try to enable path MTU discovery(neww age timer: 10 min)
RP/0/0/CPU0:Jan 8 16:09:25.214 : tcp[399]: [t21] Path mtu is ON (age-timer: 10)
RP/0/0/CPU0:Jan 8 16:19:25.233 : tcp[399]: [t21] PCB 0x12157020: Trying next higher MTU: 1240
RP/0/0/CPU0:Jan 8 16:21:25.245 : tcp[399]: [t21] PCB 0x12157020: Trying next higher MTU: 1240
RP/0/0/CPU0:Jan 8 16:23:25.256 : tcp[399]: [t21] PCB 0x12157020: Trying next higher MTU: 1240

R4 - ACTIVE - BGP Session restart - TCP SYN:

! - Once BGP session is cleared
! - TCP SYN sourced from R4 - Active
! - MSS calculation takes place and is influenced by TCP PMTUD

2734 4.810311 192.168.0.4 192.168.0.1 TCP 62 32077 179 [SYN] Seq=0 Win=16384
Len=0 **MSS=1460** WS=1

Frame 2734: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:d7:7e:f6 (fa:16:3e:d7:7e:f6), Dst: fa:16:3e:8f:8f:54
(fa:16:3e:8f:8f:54)
Internet Protocol Version 4, Src: 192.168.0.4, Dst: 192.168.0.1
Transmission Control Protocol, Src Port: 32077, Dst Port: 179, Seq: 0, Len: 0
Source Port: 32077
Destination Port: 179
[Stream index: 25]
[TCP Segment Len: 0]
Sequence number: 0 (relative sequence number)
Acknowledgment number: 0
Header Length: 28 bytes
Flags: 0x002 (SYN)
Window size value: 16384
[Calculated window size: 16384]
Checksum: 0x6398 [unverified]
[Checksum Status: Unverified]

```
Urgent pointer: 0
Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
  Maximum segment size: 1460 bytes
    Kind: Maximum Segment Size (2)
    Length: 4
    MSS Value: 1460
  Window scale: 0 (multiply by 1)
  End of Option List (EOL)
```

R1に表示されるように、パッシブ - BGPセッションの再起動 - TCP SYN、ACK。

```
! - Once BGP session is cleared
! - TCP SYN,ACK sourced from R1 - Passive
! - MSS calculation takes place and is influenced by TCP PMTUD
```

```
2735  0.003879      192.168.0.1 192.168.0.4 TCP      62      179  32077 [SYN, ACK] Seq=0 Ack=1
Win=16384 Len=0 MSS=1460 WS=1
```

```
Frame 2735: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:8f:8f:54 (fa:16:3e:8f:8f:54), Dst: fa:16:3e:d7:7e:f6
(fa:16:3e:d7:7e:f6)
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
Transmission Control Protocol, Src Port: 179, Dst Port: 32077, Seq: 0, Ack: 1, Len: 0
  Source Port: 179
  Destination Port: 32077
  [Stream index: 25]
  [TCP Segment Len: 0]
  Sequence number: 0      (relative sequence number)
  Acknowledgment number: 1      (relative ack number)
  Header Length: 28 bytes
  Flags: 0x012 (SYN, ACK)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0xbf77 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
    Maximum segment size: 1460 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1460
    Window scale: 0 (multiply by 1)
    End of Option List (EOL)
```

TCP PMTUDが有効でBGPセッションがクリアされた後の、R4で表示されるTCPセッションの詳細 (アクティブ) :

```
! - BGP session re-established
! - as seen on R4 - Active
```

```
RP/0/0/CPU0:R4#show tcp detail pcb 0x121567f4
Fri Jan  8 16:45:13.928 UTC
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan  8 16:41:49 2021

PCB 0x121567f4, SO 0x12154460, TCPCB 0x12156190, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 10
Local host: 192.168.0.4, Local port: 32077 (Local App PID: 1052958)
Foreign host: 192.168.0.1, Foreign port: 179

Current send queue size in bytes: 0 (max 24576)
```

Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	8	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	5	3	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 1254100669 snduna: 1254100983 sndnxt: 1254100983
sndmax: 1254100983 sndwnd: 32508 sndcwnd: 4380
irs: 839938559 rcvnxt: 839938873 rcvwnd: 32527 rcvadv: 839971400

SRTT: 79 ms, RTTO: 485 ms, RTV: 406 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 30 secs

State flags: none
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R4#

TCP PMTUDが有効になり、BGPセッションがクリアされた後の、R1で表示されるTCPセッションの詳細 (パッシブ)。

! - BGP session re-established
! - as seen on R1 - Passive

RP/0/0/CPU0:R1#show tcp detail pcb 0x121558cc

Fri Jan 8 16:44:59.448 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Fri Jan 8 16:41:46 2021

PCB 0x121558cc, SO 0x121556d4, TCPCB 0x121575bc, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 10

Local host: 192.168.0.1, Local port: 179 (Local App PID: 983326)
Foreign host: 192.168.0.4, Foreign port: 32077

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	6	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	6	3	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 839938559 snduna: 839938873 sndnxt: 839938873
sndmax: 839938873 sndwnd: 32527 sndcwnd: 4380
irs: 1254100669 rcvnxt: 1254100983 rcvwnd: 32508 rcvadv: 1254133491

SRTT: 76 ms, RTTO: 454 ms, RTV: 378 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 219 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R1#

PMTUD : パスセグメントのIP MTUが小さい

前のシナリオは、PMTUDが有効になっている最初のTCPセッション確立時に何が起こるかを理解するのに役立ちました。このシナリオは上に構築されており、TCP PMTUDの動作と、確立されたTCPセッションへの影響を理解するのに役立ちます。

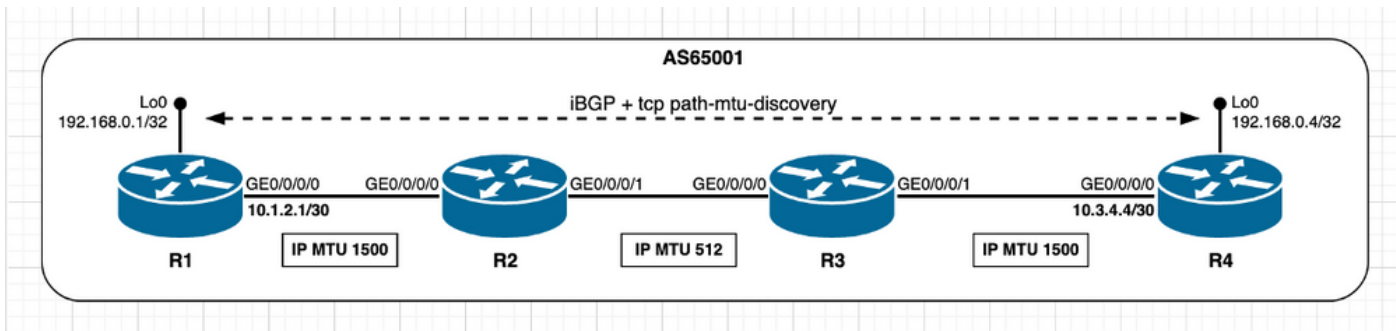


図3.2:PMTUDが有効で、パスセグメントのIP MTUが小さくなります。

前の図を参考にして、BGPセッションが確立され、R1が512バイトを超えるサイズのIPパケットによって伝送されるBGP Updateメッセージを送信するとします。PMTUDを有効にすると、DFビット(Don't Fragment)が設定されます。したがって、ノードR2はIPパケットをドロップし、ICMP(Internet Control Message Protocol)メッセージ(Destination Unreachable - type 3;フラグメンテーションが必要です (コード4))。R1に戻ります。ICMPメッセージを受信した後、ノードR1でPMTUDがトリガーされ、パスが最小のIP MTUを確立しようとしてします。これは、新しいTCPセッションMSS値と見なされる、明確に定義されたプラットフォームレベルのセットから次に低い値を使用して行われます。次に、TCPは新しいMSS値を使用して元のBGPアップデートを再送信し、このプロセスはICMPメッセージ(Destination Unreachable - type 3;フラグメンテーションが必要 - コード4)は受信されなくなりました。これは、使用中のMSS値が、送信されたすべてのパケットが最小パスセグメントIP MTUに該当することを意味します。時間が経過すると、PmtuAgerタイマーによって規定されたPMTUDが逆方向にプラットフォームレベルを通過し、MSSを最大値に戻します。任意の時点で、ICMPメッセージ(Destination Unreachable - type 3;Fragmentation needed - Code 4)が再び受信されると、PMTUDは前述のように動作します。

次の出力は、上記で説明したPMTUDの動作を示し、確立されたTCPセッションのシナリオから開始します。ここでは、Cisco IOS XRノードR4がアクティブな役割を果たしているため、TCP接続が管理され、宛先ポート179でR1とのTCPセッションが開きます。どちらのノードも、インターフェイスのデフォルトのIP MTU値を使用します。このシナリオの初期MSS計算は、次のように要約できます。

- R2ノードとR3ノード間の中間セグメントは、デフォルト以外のIP MTU 512バイトを使用します。
- R1とR4は、インターフェイスでデフォルトのMTU値を使用します。
- TCPパスMTUディスカバリーが有効になっている。
- TCPピアは直接接続されていません。R4はBGP接続を管理します。R4は1460バイトのMSSでSYNを送信します。1500 (インターフェイスIP MTU) :20(minTCP_H):20(minIP_H)R1は1460バイトのMSSでSYN、ACKを送信します。[Received MSS ;ローカルの初期MSS]です。受信MSS 1460バイト。ローカル初期MSS 1460バイト。最小のMSS値が両方のピアで使用されます。

R4から送信されたTCP SYN:

```
! - Initial TCP session establishment
! - TCP SYN sourced from R4
```

```
392 6.752774 192.168.0.4 192.168.0.1 TCP 62 32449 179 [SYN] Seq=0 Win=16384
Len=0 MSS=1460 WS=1
```

```
Frame 392: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:5c:f1:80 (fa:16:3e:5c:f1:80), Dst: fa:16:3e:42:18:05
(fa:16:3e:42:18:05)
```

```
Internet Protocol Version 4, Src: 192.168.0.4, Dst: 192.168.0.1
Transmission Control Protocol, Src Port: 32449, Dst Port: 179, Seq: 0, Len: 0
  Source Port: 32449
  Destination Port: 179
  [Stream index: 10]
  [TCP Segment Len: 0]
  Sequence number: 0 (relative sequence number)
  Acknowledgment number: 0
  Header Length: 28 bytes
  Flags: 0x002 (SYN)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0x6858 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
    Maximum segment size: 1460 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1460
    Window scale: 0 (multiply by 1)
    End of Option List (EOL)
```

TCP SYN、R1からACK:

```
! - Initial TCP session establishment
! - TCP SYN,ACK sourced from R1
```

```
393 0.003628 192.168.0.1 192.168.0.4 TCP 62 179 32449 [SYN, ACK] Seq=0 Ack=1
Win=16384 Len=0 MSS=1460 WS=1
```

```
Frame 393: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
(fa:16:3e:5c:f1:80)
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
Transmission Control Protocol, Src Port: 179, Dst Port: 32449, Seq: 0, Ack: 1, Len: 0
  Source Port: 179
  Destination Port: 32449
  [Stream index: 10]
  [TCP Segment Len: 0]
  Sequence number: 0 (relative sequence number)
  Acknowledgment number: 1 (relative ack number)
  Header Length: 28 bytes
  Flags: 0x012 (SYN, ACK)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0x509e [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (8 bytes), Maximum segment size, Window scale, End of Option List (EOL)
    Maximum segment size: 1460 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1460
    Window scale: 0 (multiply by 1)
    End of Option List (EOL)
```

BGPセッションが確立されると、ノードR1はBGP Updateメッセージを送信し、ICMPメッセージ (Destination Unreachable - type 3;フラグメンテーションが必要 - コード4)をノードR2から返しません。

これは、BGPアップデートメッセージを伝送するIPパケットにDFビットが設定されており、R2/R3セグメントで使用されるIP MTUが1116バイトのIPパケットサイズより小さいためです。前

述したように、ICMPメッセージの受信によってPMTUDがトリガーされます。

R1 ICMPで、タイプ3/コード4メッセージが受信されます。

```
! - as seen from R1 - Passive
! - After session is established R1 sends BGP Update message with IP length of 1116 Bytes
! - note IP Header Flags shows DF bit set

528      5.893055      192.168.0.1 192.168.0.4 BGP      1130      UPDATE Message, KEEPALIVE Message

Frame 528: 1130 bytes on wire (9040 bits), 1130 bytes captured (9040 bits) on interface 0
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
(fa:16:3e:5c:f1:80)
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
  Total Length: 1116
  Identification: 0x8c37 (35895)
  Flags: 0x02 (Don't Fragment)
  Fragment offset: 0
  Time to live: 255
  Protocol: TCP (6)
  Header checksum: 0xe09a [validation disabled]
  [Header checksum status: Unverified]
  Source: 192.168.0.1
  Destination: 192.168.0.4
  [Source GeoIP: Unknown]
  [Destination GeoIP: Unknown]
Transmission Control Protocol, Src Port: 179, Dst Port: 32449, Seq: 318, Ack: 251, Len: 1076
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - KEEPALIVE Message
<snip>
```

```
! - as seen from R1 - Passive
! - IP MTU on R2/R3 is lower than IP packet length and DF bit is set
! - R1 receives ICMP error message from R2
! - note R2 ICMP error message carries Next-Hop MTU
! - "The size in octets of the largest datagram that could be forwarded, along the path of
!   the original datagram, without being fragmented at this router. The size includes the
!   IP header and IP data, and does not include any lower-level headers."
```

```
529      0.002423      10.2.3.1      192.168.0.1 ICMP      110      Destination unreachable
(Fragmentation needed)
```

```
Frame 529: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface 0
Ethernet II, Src: fa:16:3e:5c:f1:80 (fa:16:3e:5c:f1:80), Dst: fa:16:3e:42:18:05
(fa:16:3e:42:18:05)
Internet Protocol Version 4, Src: 10.2.3.1, Dst: 192.168.0.1
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
  Total Length: 96
  Identification: 0x0001 (1)
  Flags: 0x00
  Fragment offset: 0
  Time to live: 255
  Protocol: ICMP (1)
  Header checksum: 0xac97 [validation disabled]
  [Header checksum status: Unverified]
  Source: 10.2.3.1
  Destination: 192.168.0.1
```



```

[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Internet Control Message Protocol
  Type: 3 (Destination unreachable)
  Code: 4 (Fragmentation needed)
Checksum: 0x2d52 [correct]
[Checksum Status: Good]
Length: 17
[Length of original datagram: 68]
Unused: 0011
  MTU of next hop: 512
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
Total Length: 1116
Identification: 0x8c37 (35895)
Flags: 0x02 (Don't Fragment)
Fragment offset: 0
Time to live: 254
Protocol: TCP (6)
Header checksum: 0xe19a [validation disabled]
[Header checksum status: Unverified]
Source: 192.168.0.1
Destination: 192.168.0.4
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Transmission Control Protocol, Src Port: 179, Dst Port: 32449, Seq: 2847698730, Ack:
2130367817
Border Gateway Protocol - UPDATE Message
[Packet size limited during capture: IPv4 truncated]

```

ノードR1では、ICMPメッセージによってトリガーされ、TCP PMTUDは、明確に定義されたプラトー(IP MTU)レベルの次に低い値を使用して、エンドツーエンドの最小IP MTUを確立しようとします。これらのプラトーのレベルは、[RFC1191 - Path MTU Discovery](#)に記載されています。

```

MTU plateaus from RFC 1191
- values include both TCP and IP headers
65535
32000
17914
8166
4352
2002
1492
1006
508
296
68

```

しかし、ICMP(Destination Unreachable - type 3;Fragmentation needed - Code 4)メッセージをノードR1が受信し、次に示すようにネクストホップのMTUを伝送し、ノードR1はこの値を使用します。この例では512バイトで、TCPセッションMSS値を調整します。元のTCPセグメントの長さが1076バイトであることに注意してください。したがって、元のTCPセグメントを再送信するには3つのパケットが必要です。

R1 - PASSIVE - PMTUD動作で確認できます。

```

! - As seen from R1 - Passive
! - Hint is provided by ICMP unreachable message MTU of next-hop field: 512 bytes
! - R1 then considers this value and retransmits BGP Update split in three distinct packets

```

! - Sum of TCP length = 472 + 472 + 132 = 1076 bytes

530 0.007497 192.168.0.1 192.168.0.4 TCP 526 [TCP Out-Of-Order] 179 32449 [ACK]
Seq=318 Ack=251 Win=32593 Len=472

532 0.015374 192.168.0.1 192.168.0.4 TCP 526 [TCP Retransmission] 179 32449
[ACK] Seq=790 Ack=251 Win=32593 Len=472

533 0.004129 192.168.0.1 192.168.0.4 TCP 186 [TCP Retransmission] 179 32449
[PSH, ACK] Seq=1262 Ack=251 Win=32593 Len=132

前述のように、すべてのパケットが時間の経過とともに送信されると、PMTUDはPmtuAgerタイマーによって規定される逆方向のプラタレベルを通過し、設定されているシナリオに従ってMSSを最大値まで上げようとします。

R1 - PMTUDで定義されたプレート全体で確認できます。

! - As seen from R1 - Passive - 'debug tcp pmtud' and 'debug icmp' active
! - TCP PMTUD is triggered once ICMP unreachable received

RP/0/0/CPU0:May 12 09:09:22.763 UTC: ipv4_io[266]: IPv4 ICMP: Received ICMP too big from
192.168.0.1 about 192.168.0.4, MTU=512

RP/0/0/CPU0:May 12 09:09:22.763 UTC: ipv4_io[266]: ipv4_icmp_unreachable_rcvd ICMP unreach
recvd: sending pak(0xb0c07d8f) to transport: 6, tid: 5

RP/0/0/CPU0:May 12 09:09:22.763 UTC: ipv4_io[266]: ip_icmp_lib_ipv4_receive: sending
pak(0xb0c07d8f) to transport: 1, tid: 5

RP/0/0/CPU0:May 12 09:09:22.763 UTC: tcp[399]: [t4] PCB 0x15393770: Process ICMP Dest-unreach
(next hop mtu: 512)

! - attempt new MSS 472 = MTU of next-hop(512) - TCP_H(20) - IP_H(20)

RP/0/0/CPU0:May 12 09:09:22.763 UTC: tcp[399]: [t4] PCB 0x15393770: Process ICMP Dest-unreach
(next hop mtu: 512)

RP/0/0/CPU0:May 12 09:09:22.763 UTC: tcp[399]: [t4] PCB 0x15393770: Try to use new MSS: 472

RP/0/0/CPU0:May 12 09:09:22.763 UTC: tcp[399]: [t4] PCB 0x15393770, New path MTU decided to use:
472 configured tp_user_mss 0

! - over time PMTUD attempts to raise MSS as per egress interface configured MTU

RP/0/0/CPU0:May 12 09:19:22.782 UTC: tcp[399]: [t23] PCB 0x15393770: Trying next higher MTU: 966

RP/0/0/CPU0:May 12 09:21:22.793 UTC: tcp[399]: [t23] PCB 0x15393770: Trying next higher MTU:
1452

RP/0/0/CPU0:May 12 09:23:22.805 UTC: tcp[399]: [t23] PCB 0x15393770: Trying next higher MTU:
1460

これらの出力で最終的な状態を確認できます。特に、ノードR1が示す最小および最大のMSS値に注意してください。この値は、PMTUDがトリガーされたことを示し、信号を強調表示します。

R4に表示されるTCPセッションの詳細 - ACTIVE:

! - Final stage as seen from R4 - Active

RP/0/0/CPU0:R4#show tcp detail pcb 0x153913b8

Wed May 12 10:09:43.246 UTC

=====

Connection state is ESTAB, I/O status: 0, socket status: 0

Established at Wed May 12 09:02:07 2021

PCB 0x153913b8, SO 0x153917f0, TCPCB 0x1538fb58, vrfid 0x60000000,

Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 382

Local host: 192.168.0.4, Local port: 32449 (Local App PID: 1196319)

Foreign host: 192.168.0.1, Foreign port: 179

(Local App PID/instance/SPL_APP_ID: 1196319/1/0)

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	72	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	71	69	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 2130367566 snduna: 2130368957 sndnxt: 2130368957
sndmax: 2130368957 sndwnd: 31453 sndcwnd: 2920
irs: 2847698412 rcvnxt: 2847700946 rcvwnd: 31799 rcvadv: 2847732745

SRTT: 220 ms, RTTO: 300 ms, RTV: 12 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 10, connect retry interval: 30 secs

State flags: none
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0
Socket misc info : Rcv data size (sb_cc) 0, so_qlen 0,
so_q0len 0, so_qlimit 0, so_error 0
so_auto_rearm 1

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:
Num of peers with authentication info: 0

RP/0/0/CPU0:R4#

R1に表示されるTCPセッションの詳細: パッシブ:

! - Final stage as seen from R1 - Passive

RP/0/0/CPU0:R1#show tcp detail pcb 0x15393770
Wed May 12 10:12:41.432 UTC

=====
Connection state is ESTAB, I/O status: 240, socket status: 0

Established at Wed May 12 09:02:05 2021

PCB 0x15393770, SO 0x15394ea0, TCPCB 0x15391c0c, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 382
Local host: 192.168.0.1, Local port: 179 (Local App PID: 1192224)
Foreign host: 192.168.0.4, Foreign port: 32449
(Local App PID/instance/SPL_APP_ID: 1192224/1/0)

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	75	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	73	71	0
KeepAlive	1	0	0
PmtuAger	28	27	41595
GiveUp	0	0	0
Throttle	0	0	0

iss: 2847698412 snduna: 2847701003 sndnxt: 2847701003
sndmax: 2847701003 sndwnd: 31742 sndcwnd: 4380
irs: 2130367566 rcvnxt: 2130369014 rcvwnd: 31396 rcvadp: 2130400410

SRTT: 224 ms, RTTO: 300 ms, RTV: 23 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 259 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: PMTU ager
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 472, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0
Socket misc info : Rcv data size (sb_cc) 0, so_qlen 0,
so_q0len 0, so_qlimit 0, so_error 0
so_auto_rearm 1

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x20 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:
Num of peers with authentication info: 0

RP/0/0/CPU0:R1#

最後に、任意の時点でICMP(Destination Unreachable - type 3 (宛先到達不能 - タイプ3)、
Fragmentation needed - Code 4)メッセージが再び受信された後、PMTUDが前述のように再度動

作します。

R1 - PASSIVE - PMTUDが再びトリガーされました。

! - As seen from R1 - Passive
! - TCP PMTUD is again triggered upon new ICMP unreachable received
! - Behavior can be triggered via clearing redistributed, network and aggregate routes originated

```
RP/0/0/CPU0:R1#clear bgp ipv4 all self-originated
Wed May 12 10:19:06.836 UTC
RP/0/0/CPU0:R1#
```

! - New BGP update message is sourced from R1 after clear bgp command

```
1707 1.712657 192.168.0.1 192.168.0.4 BGP 1121 UPDATE Message
```

Frame 1707: 1121 bytes on wire (8968 bits), 1121 bytes captured (8968 bits) on interface 0
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
(fa:16:3e:5c:f1:80)

Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4

0100 = Version: 4

.... 0101 = Header Length: 20 bytes (5)

Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)

Total Length: 1107

Identification: 0x1a38 (6712)

Flags: 0x02 (Don't Fragment)

Fragment offset: 0

Time to live: 255

Protocol: TCP (6)

Header checksum: 0x52a3 [validation disabled]

[Header checksum status: Unverified]

Source: 192.168.0.1

Destination: 192.168.0.4

[Source GeoIP: Unknown]

[Destination GeoIP: Unknown]

Transmission Control Protocol, Src Port: 179, Dst Port: 32449, Seq: 2705, Ack: 1562, Len: 1067
Border Gateway Protocol - UPDATE Message

! - ICMP Destination Unreachable / Fragmentation needed is received and triggers PMTUD

```
1708 0.001614 10.2.3.1 192.168.0.1 ICMP 110 Destination unreachable  
(Fragmentation needed)
```

Frame 1708: 110 bytes on wire (880 bits), 110 bytes captured (880 bits) on interface 0
Ethernet II, Src: fa:16:3e:5c:f1:80 (fa:16:3e:5c:f1:80), Dst: fa:16:3e:42:18:05
(fa:16:3e:42:18:05)

Internet Protocol Version 4, Src: 10.2.3.1, Dst: 192.168.0.1

0100 = Version: 4

.... 0101 = Header Length: 20 bytes (5)

Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)

Total Length: 96

Identification: 0x0002 (2)

Flags: 0x00

Fragment offset: 0

Time to live: 255

Protocol: ICMP (1)

Header checksum: 0xac96 [validation disabled]

[Header checksum status: Unverified]

Source: 10.2.3.1

```

Destination: 192.168.0.1
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Internet Control Message Protocol
  Type: 3 (Destination unreachable)
  Code: 4 (Fragmentation needed)
Checksum: 0x3b73 [correct]
[Checksum Status: Good]
Length: 17
[Length of original datagram: 68]
Unused: 0011
MTU of next hop: 512
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
  Total Length: 1107
  Identification: 0x1a38 (6712)
  Flags: 0x02 (Don't Fragment)
  Fragment offset: 0
  Time to live: 254
  Protocol: TCP (6)
  Header checksum: 0x53a3 [validation disabled]
  [Header checksum status: Unverified]
  Source: 192.168.0.1
  Destination: 192.168.0.4
  [Source GeoIP: Unknown]
  [Destination GeoIP: Unknown]
  Transmission Control Protocol, Src Port: 179, Dst Port: 32449, Seq: 2847701117, Ack:
2130369128
  Border Gateway Protocol - UPDATE Message

```

```

! - Note new/updated MSS value and PmtuAger
! - MSS 472 ; Aligned with "MTU of next hop" value contained in ICMP message

```

```

RP/0/0/CPU0:R1#show tcp detail pcb 0x15393770
Wed May 12 10:19:31.494 UTC

```

```

=====
Connection state is ESTAB, I/O status: 240, socket status: 0
Established at Wed May 12 09:02:05 2021

```

```

PCB 0x15393770, SO 0x15394ea0, TCPCB 0x15391c0c, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 382
Local host: 192.168.0.1, Local port: 179 (Local App PID: 1192224)
Foreign host: 192.168.0.4, Foreign port: 32449
(Local App PID/instance/SPL_APP_ID: 1192224/1/0)

```

```

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768)  mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

```

Timer	Starts	Wakeups	Next(msec)
Retrans	83	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	80	77	0
KeepAlive	1	0	0
PmtuAger	32	30	575401
GiveUp	0	0	0
Throttle	0	0	0

```

  iss: 2847698412  snduna: 2847702184  sndnxt: 2847702184
sndmax: 2847702184  sndwnd: 32173  sndcwnd: 944
  irs: 2130367566  rcvnxt: 2130369147  rcvwnd: 32730  rcvadv: 2130401877

```



```
RP/0/0/CPU0:Feb 23 17:05:37.239 : tcp[399]: [t27] PCB 0x12152adc: Trying next lower MTU: 966
<<<<<<<< HERE: Plateau 1006
```

```
RP/0/0/CPU0:Feb 23 17:05:38.109 : tcp[399]: [t4] PCB 0x12152adc: Process ICMP Dest-unreach (next
hop mtu: 33554432)
```

```
RP/0/0/CPU0:Feb 23 17:05:38.109 : tcp[399]: [t4] PCB 0x12152adc: Invalid next hop mtu
(33554432), ignore it
```

```
RP/0/0/CPU0:Feb 23 17:05:39.829 : tcp[399]: [t27] PCB 0x12152adc: Trying next lower MTU: 468
<<<<<<<< HERE: Plateau 508
```

次のステップとして、すべてのインターフェイスでLabel Distribution Protocol(LDP)を使用する同じシナリオを検討します。ここでの目標は、MPLSが有効な環境で以前のシナリオと観察できる違いを理解することです。

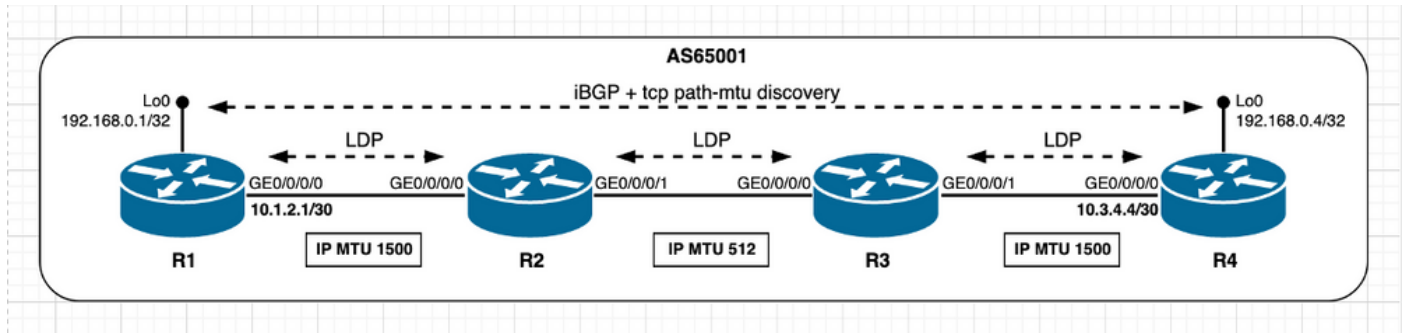


図3.3:PMTUDが有効で、パスセグメントのIP MTU - MPLSシナリオが低い

まず、次に示すように、PMTUDトリガーの前に確立されたBGPセッションの初期段階を考慮します。

TCP(BGP)の初期状態 (R4 - ACTIVE - MPLS対応シナリオ) :

```
! - as seen on R4 - Active
! - TCP path MTU discovery enabled
! - MPLS LDP enabled
! - TCP session initial state
```

```
RP/0/0/CPU0:R4#show tcp detail pcb 0x153bdaf0
Mon May 17 08:32:16.673 UTC
```

```
=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Mon May 17 08:31:57 2021
```

```
PCB 0x153bdaf0, SO 0x153acc80, TCPCB 0x153acea8, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 757
Local host: 192.168.0.4, Local port: 57400 (Local App PID: 1196319)
Foreign host: 192.168.0.1, Foreign port: 179
(Local App PID/instance/SPL_APP_ID: 1196319/1/0)
```

```
Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)
```

Timer	Starts	Wakeup	Next(msec)
Retrans	5	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	2	1	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0

Throttle 0 0 0

iss: 1386459919 snduna: 1386460037 sndnxt: 1386460037
sndmax: 1386460037 sndwnd: 32726 sndcwnd: 4380
irs: 3874414679 rcvnxt: 3874414864 rcvwnd: 32678 rcvadp: 3874447542

SRTT: 48 ms, RTTO: 300 ms, RTV: 228 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 10, connect retry interval: 30 secs

State flags: none
Feature flags: Win Scale, Nagle, Path MTU
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0
Socket misc info : Rcv data size (sb_cc) 0, so_qlen 0,
so_q0len 0, so_qlimit 0, so_error 0
so_auto_rearm 1

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 1 Label Stack: 0x5dc2
Num of peers with authentication info: 0

RP/0/0/CPU0:R4#

TCP(BGP)の初期状態 (R1 : パッシブ - MPLS対応シナリオ) :

! - as seen on R1 - Passive
! - TCP path MTU discovery enabled
! - MPLS LDP enabled
! - TCP session initial state

RP/0/0/CPU0:R1#show tcp detail pcb 0x153acc8c
Mon May 17 08:32:56.618 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Mon May 17 08:31:55 2021

PCB 0x153acc8c, SO 0x153adad4, TCPCB 0x153adcfc, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 757
Local host: 192.168.0.1, Local port: 179 (Local App PID: 1192224)
Foreign host: 192.168.0.4, Foreign port: 57400
(Local App PID/instance/SPL_APP_ID: 1192224/1/0)

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes

Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	3	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	3	1	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 3874414679 snduna: 3874414864 sndnxt: 3874414864
sndmax: 3874414864 sndwnd: 32678 sndcwnd: 4380
irs: 1386459919 rcvnxt: 1386460037 rcvwnd: 32726 rcvadv: 1386492763

SRTT: 45 ms, RTTO: 300 ms, RTV: 239 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 229 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 1460, peer MSS 1460, min MSS 1460, max MSS 1460

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0
Socket misc info : Rcv data size (sb_cc) 0, so_qlen 0,
so_q0len 0, so_qlimit 0, so_error 0
so_auto_rearm 1

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x20 PD ctx: size: 0 data:
Num Labels: 1 Label Stack: 0x5dc3
Num of peers with authentication info: 0

RP/0/0/CPU0:R1#

このMPLS対応のシナリオでは、TCP(LDP)セッションの詳細が確立されていることが確認されま
す。TCP(BGP)セッションのMSS計算に関する前述の説明はすべて、TCP(LDP)セッションにも適
用されることに注意してください。たとえば、ノードR3とR2 TCP(LDP)セッションMSSの計算
を次のように要約できます。

- R2とR3の両方が、デフォルト以外の512バイトのIP MTUを使用します。
- パスMTUディスカバリが有効になっている。
- TCPピアが直接接続されていない (ループバックインターフェイス間でTCPセッションが確
立される)。 R3はLDP接続を管理します。 R3は472バイトのMSSでSYNを送信します。
512 (インターフェイスIP MTU) :20(minTCP_H):20(minIP_H)R2は472バイトのMSSで

SYN、ACKを送信します。[Received MSS;ローカルの初期MSS]です。受信MSS 472バイト。
ローカル初期MSS 472バイト。最小のMSS値が両方のピアで使用されます。

R3 - ACTIVE - MPLS対応シナリオに表示されるTCP(LDP)セッションの詳細：

! - as seen on R3 - Active
! - TCP path MTU discovery enabled
! - MPLS LDP enabled
! - TCP session initial state

RP/0/0/CPU0:R3#show tcp detail pcb 0x15393fbc
Mon May 17 08:33:30.627 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Mon May 17 08:30:04 2021

PCB 0x15393fbc, SO 0x15393d94, TCPCB 0x153941b4, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 970
Local host: 192.168.0.3, Local port: 57146 (Local App PID: 1151216)
Foreign host: 192.168.0.2, Foreign port: 646
(Local App PID/instance/SPL_APP_ID: 1151216/0/0)

Current send queue size in bytes: 0 (max 16384)
Current receive queue size in bytes: 0 (max 16384) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 60)

Timer	Starts	Wakeups	Next(msec)
Retrans	8	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	6	4	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 2917752466 snduna: 2917752838 sndnxt: 2917752838
sndmax: 2917752838 sndwnd: 16013 sndcwnd: 944
irs: 228184383 rcvnxt: 228184763 rcvwnd: 16005 rcvadv: 228200768

SRTT: 103 ms, RTTO: 580 ms, RTV: 477 ms, KRTT: 0 ms
minRTT: 9 ms, maxRTT: 279 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 1, connect retry interval: 3 secs

State flags: none
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 472, peer MSS 472, min MSS 472, max MSS 472

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_SEL, SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP

Socket receive buffer: Low/High watermark 1/16384
Socket send buffer : Low/High watermark 2048/16384, Notify threshold 0
Socket misc info : Rcv data size (sb_cc) 0, so_qlen 0,
so_q0len 0, so_qlimit 0, so_error 0
so_auto_rearm 1

PDU information:

#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 1 Label Stack: 0x5dc2
Num of peers with authentication info: 0

RP/0/0/CPU0:R3#

R2 (パッシブ - MPLS対応) のシナリオに表示されるTCP(LDP)セッションの詳細 :

! - as seen on R2 - Passive
! - TCP path MTU discovery enabled
! - MPLS LDP enabled
! - TCP session initial state

RP/0/0/CPU0:R2#show tcp detail pcb 0x153a1f44
Mon May 17 08:34:28.843 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Mon May 17 08:30:31 2021

PCB 0x153a1f44, SO 0x153a1d1c, TCPCB 0x153a213c, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 970
Local host: 192.168.0.2, Local port: 646 (Local App PID: 1151216)
Foreign host: 192.168.0.3, Foreign port: 57146
(Local App PID/instance/SPL_APP_ID: 1151216/0/0)

Current send queue size in bytes: 0 (max 16384)
Current receive queue size in bytes: 0 (max 16384) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 60)

Timer	Starts	Wakeups	Next(msec)
Retrans	7	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	7	5	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 228184383 snduna: 228184763 sndnxt: 228184763
sndmax: 228184763 sndwnd: 16005 sndcwnd: 944
irs: 2917752466 rcvnxt: 2917752856 rcvwnd: 15995 rcvadv: 2917768851

SRTT: 95 ms, RTTO: 561 ms, RTV: 466 ms, KRTT: 0 ms
minRTT: 0 ms, maxRTT: 219 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: Win Scale, Nagle, **Path MTU**
Request flags: Win Scale

Datagrams (in bytes): MSS 472, peer MSS 472, min MSS 472, max MSS 472

```
Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none
```

```
Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_SEL, SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/16384
Socket send buffer : Low/High watermark 2048/16384, Notify threshold 0
Socket misc info : Rcv data size (sb_cc) 0, so_qlen 0,
                  so_q0len 0, so_qlimit 0, so_error 0
                  so_auto_rearm 1
```

```
PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x60 PD ctx: size: 0 data:
Num Labels: 1 Label Stack: 0x5dc1
Num of peers with authentication info: 0
```

```
RP/0/0/CPU0:R2#
```

BGPセッションが確立されると、R1はBGP Updateメッセージを送信し、ICMPメッセージ (Destination Unreachable - type 3; Fragmentation needed - コード4)は、ノードR1でTCP PMTUDをトリガーするノードR2からのリターンで発生します。これは、BGPアップデートメッセージを伝送するIPパケットのDFビットが設定され、R2/R3セグメントで使用されるIPのMTUが5111111111116バイト。以前と同様に、このICMPメッセージの受信によってPMTUDがトリガーされます。MPLS対応シナリオと以前の非MPLSシナリオの違いは、ノードR2 ICMPメッセージに含まれるネクストホップ値のMTU(宛先到達不能 - タイプ3; フラグメンテーションが必要 - コード4)。このMPLS対応のシナリオでは、ネクストホップ値のMTUが4バイトの追加MPLSオーバーヘッドを占めて、次の出力に示すように、R2の出力MPLSラベルスタックを占めていることを意味します。

R1で見られるTCPパスMTUディスカバリ : パッシブ - MPLS対応シナリオ :

```
! - as seen from R1 - Passive
! - R1 sends BGP Update message with IP length of 1116 Bytes
! - Note MPLS Header as packet is to be label-switched (single label ; IGP label)
! - note IP Header Flags shows DF bit set

455      0.044859      192.168.0.1 192.168.0.4 BGP      1134      UPDATE Message, KEEPALIVE Message

Frame 455: 1134 bytes on wire (9072 bits), 1134 bytes captured (9072 bits) on interface 0
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
(fa:16:3e:5c:f1:80)
MultiProtocol Label Switching Header, Label: 24002, Exp: 6, S: 1, TTL: 255
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
Total Length: 1116
  Identification: 0xc6dd (50909)
  Flags: 0x02 (Don't Fragment)
    0... .... = Reserved bit: Not set
    .1.. .... = Don't fragment: Set
    ..0. .... = More fragments: Not set
  Fragment offset: 0
  Time to live: 255
  Protocol: TCP (6)
```

```
Header checksum: 0xa5f4 [validation disabled]
[Header checksum status: Unverified]
Source: 192.168.0.1
Destination: 192.168.0.4
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Transmission Control Protocol, Src Port: 179, Dst Port: 57400, Seq: 242, Ack: 175, Len: 1076
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - KEEPALIVE Message
<snip>
```

```
! - as seen from R1 - Passive
! - IP MTU on R2/R3 of 512 bytes is lower than IP packet length and DF bit is set
! - R1 receives ICMP error message from R2
! - note R2 ICMP error message carries Next-Hop MTU
! - "The size in octets of the largest datagram that could be forwarded, along the path of
!   the original datagram, without being fragmented at this router. The size includes the
!   IP header and IP data, and does not include any lower-level headers."
! - In present MPLS-enabled scenario Next-Hop MTU value is 508 bytes
! - In previous non-MPLS scenario Next-Hop MTU value was 512 bytes
```

```
456      0.014117      10.2.3.1      192.168.0.1 ICMP      182      Destination unreachable
(Fragmentation needed)
```

```
Frame 456: 182 bytes on wire (1456 bits), 182 bytes captured (1456 bits) on interface 0
Ethernet II, Src: fa:16:3e:5c:f1:80 (fa:16:3e:5c:f1:80), Dst: fa:16:3e:42:18:05
(fa:16:3e:42:18:05)
```

```
Internet Protocol Version 4, Src: 10.2.3.1, Dst: 192.168.0.1
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
  Total Length: 168
  Identification: 0x001f (31)
  Flags: 0x00
    0... .... = Reserved bit: Not set
    .0.. .... = Don't fragment: Not set
    ..0. .... = More fragments: Not se
```

```
Fragment offset: 0
```

```
Time to live: 251
```

```
Protocol: ICMP (1)
```

```
Header checksum: 0xb031 [validation disabled]
```

```
[Header checksum status: Unverified]
```

```
Source: 10.2.3.1
```

```
Destination: 192.168.0.1
```

```
[Source GeoIP: Unknown]
```

```
[Destination GeoIP: Unknown]
```

```
Internet Control Message Protocol
```

```
Type: 3 (Destination unreachable)
```

```
Code: 4 (Fragmentation needed)
```

```
Checksum: 0x5199 [correct]
```

```
[Checksum Status: Good]
```

```
Length: 17
```

```
[Length of original datagram: 68]
```

```
Unused: 0011
```

```
MTU of next hop: 508
```

```
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
```

```
Transmission Control Protocol, Src Port: 179, Dst Port: 57400, Seq: 3874414921, Ack:
```

```
1386460094
```

```
Border Gateway Protocol - UPDATE Message
```

```
! - As seen from R1 - Passive
! - Hint is provided by ICMP unreachable message MTU of next-hop field: 508 bytes
! - R1 then considers this value and retransmits BGP Update split in three distinct packets
! - Sum of TCP length = 468 + 468 + 140 = 1076 bytes
```

```
457      0.006689      192.168.0.1 192.168.0.4 TCP      526      [TCP Retransmission] 179  57400
[ACK] Seq=242 Ack=175 Win=32669 Len=468
460      0.004001      192.168.0.1 192.168.0.4 TCP      526      [TCP Retransmission] 179  57400
[ACK] Seq=710 Ack=175 Win=32669 Len=468
461      0.001788      192.168.0.1 192.168.0.4 TCP      198      [TCP Retransmission] 179  57400
[PSH, ACK] Seq=1178 Ack=175 Win=32669 Len=140
463      0.056695      192.168.0.4 192.168.0.1 TCP      54      57400 179 [ACK] Seq=175 Ack=1318
Win=31545 Len=0
```

```
! - As seen from R1 - Passive - 'debug tcp pmtud' and 'debug icmp' active
! - TCP PMTUD is triggered once ICMP unreachable received
```

```
RP/0/0/CPU0:May 17 08:29:56.131 UTC: tcp[399]: [t1] Try to enable path MTU discovery(neww age
timer: 10 min)
RP/0/0/CPU0:May 17 08:29:56.131 UTC: tcp[399]: [t1] Path mtu is ON (age-timer: 10)
RP/0/0/CPU0:May 17 08:35:51.726 UTC: ipv4_io[266]: ip_icmp_lib_ipv4_receive: Receiving
pak(0xb0c07d8f) tid: 5
RP/0/0/CPU0:May 17 08:35:51.726 UTC: ipv4_io[266]: Entering ipv4_mtu_update_cb
RP/0/0/CPU0:May 17 08:35:51.726 UTC: ipv4_io[266]: IPv4 ICMP: Received ICMP too big from
192.168.0.1 about 192.168.0.4, MTU=508
RP/0/0/CPU0:May 17 08:35:51.726 UTC: ipv4_io[266]: ipv4_icmp_unreachable_rcvd ICMP unreach
recvd: sending pak(0xb0c07d8f) to transport: 6, tid: 5
RP/0/0/CPU0:May 17 08:35:51.726 UTC: ipv4_io[266]: ip_icmp_lib_ipv4_receive: sending
pak(0xb0c07d8f) to transport: 1, tid: 5
RP/0/0/CPU0:May 17 08:35:51.726 UTC: tcp[399]: [t4] PCB 0x153acc8c: Process ICMP Dest-unreach
(next hop mtu: 508)
```

```
! - attempt new MSS 468 = MTU of next-hop(508) - TCP_H(20) - IP_H(20)
```

```
RP/0/0/CPU0:May 17 08:35:51.726 UTC: tcp[399]: [t4] PCB 0x153acc8c: Try to use new MSS: 468
RP/0/0/CPU0:May 17 08:35:51.726 UTC: tcp[399]: [t4] PCB 0x153acc8c, New path MTU decided to use:
468 configured tp_user_mss 0
```

```
! - over time PMTUD attempts to raise MSS as per egress interface configured MTU
```

```
RP/0/0/CPU0:May 17 08:45:51.745 UTC: tcp[399]: [t29] PCB 0x153acc8c: Trying next higher MTU: 966
RP/0/0/CPU0:May 17 08:47:51.757 UTC: tcp[399]: [t29] PCB 0x153acc8c: Trying next higher MTU:
1452
RP/0/0/CPU0:May 17 08:49:51.769 UTC: tcp[399]: [t29] PCB 0x153acc8c: Trying next higher MTU:
1460
```

R1 - PASSIVE - TCP PMTUD triggered - MPLS対応のシナリオで確認できます。

```
! - as seen on R1 - Passive
! - R1 session details after TCP PMTUD trigger
```

```
RP/0/0/CPU0:R1#show tcp detail pcb 0x153acc8c
Mon May 17 08:43:07.077 UTC
=====
Connection state is ESTAB, I/O status: 240, socket status: 0
Established at Mon May 17 08:31:55 2021
```

```
PCB 0x153acc8c, SO 0x153adad4, TCPCB 0x153adcfc, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 757
Local host: 192.168.0.1, Local port: 179 (Local App PID: 1192224)
Foreign host: 192.168.0.4, Foreign port: 57400
(Local App PID/instance/SPL_APP_ID: 1192224/1/0)
```

```
Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)
```

Timer	Starts	Wakeups	Next(msec)
Retrans	15	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	14	9	0
KeepAlive	1	0	0
PmtuAger	1	0	164599
GiveUp	0	0	0
Throttle	0	0	0

```

iss: 3874414679  snduna: 3874416130  sndnxt: 3874416130
sndmax: 3874416130  sndwnd: 31412  sndcwnd: 936
irs: 1386459919  rcvnxt: 1386460246  rcvwnd: 32517  rcvadv: 1386492763

```

```

SRTT: 180 ms,  RTTO: 509 ms,  RTV: 329 ms,  KRRT: 0 ms
minRTT: 19 ms,  maxRTT: 239 ms

```

```

ACK hold time: 200 ms,  Keepalive time: 0 sec,  SYN waittime: 30 sec
Giveup time: 0 ms,  Retransmission retries: 0,  Retransmit forever: FALSE
Connect retries remaining: 0,  connect retry interval: 0 secs

```

```

State flags: PMTU ager
Feature flags: Win Scale, Nagle, Path MTU
Request flags: Win Scale

```

Datagrams (in bytes): MSS 468, peer MSS 1460, min MSS 468, max MSS 1460

```

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

```

```

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer   : Low/High watermark 2048/24576, Notify threshold 0
Socket misc info     : Rcv data size (sb_cc) 0, so_qlen 0,
                      so_q0len 0, so_qlimit 0, so_error 0
                      so_auto_rearm 1

```

```

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x20  PD ctx: size: 0  data:
Num Labels: 1  Label Stack: 0x5dc3
Num of peers with authentication info: 0

```

RP/0/0/CPU0:R1#

MPLS対応のシナリオでは、ノードR2 ICMPメッセージに含まれるネクストホップのMTUの値が出力MPLSラベルスタックの値となることに注意してください。この側面をさらに強化するには、次の例を検討します。R2でフィルタリングされたIPパケットがL3VPNサービスに関連付けられている場合は、イーサネットフレームが2つのラベル（IGPラベルとVPNラベル）を搬送していることを意味します。次に、ネクストホップのMTUに必要なラベルスタックサイズが反映されます。次の出力を参照してください。

R1 - PASSIVE - L3 VPNサービスパケットに表示されます。

```

! - as seen from R1 - Passive
! - L3 VPN service packet is sourced by node R1 and destined to node R4
! - Note presence of MPLS label stack - both IGP and VPN label are present

```


! - Note IP Total Length of 610 bytes higher than the IP MTU on R2/R3 segment
! - note IP Header Flags shows DF bit set

2024 0.302370 10.1.14.1 10.1.14.14 TELNET 632 Telnet Data ...

Frame 2024: 632 bytes on wire (5056 bits), 632 bytes captured (5056 bits) on interface 0
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
(fa:16:3e:5c:f1:80)

MultiProtocol Label Switching Header, Label: 24002, Exp: 0, S: 0, TTL: 255

0000 0101 1101 1100 0010 = MPLS Label: 24002
..... 000. = MPLS Experimental Bits: 0
..... 0 = MPLS Bottom Of Label Stack: 0
..... 1111 1111 = MPLS TTL: 255

MultiProtocol Label Switching Header, Label: 24005, Exp: 0, S: 1, TTL: 255

0000 0101 1101 1100 0101 = MPLS Label: 24005
..... 000. = MPLS Experimental Bits: 0
..... 1 = MPLS Bottom Of Label Stack: 1
..... 1111 1111 = MPLS TTL: 255

Internet Protocol Version 4, Src: 10.1.14.1, Dst: 10.1.14.14

0100 = Version: 4
.... 0101 = Header Length: 20 bytes (5)
Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
Total Length: 610
Identification: 0x7c9f (31903)
Flags: 0x02 (Don't Fragment)
0... = Reserved bit: Not set
.1.. = Don't fragment: Set
..0. = More fragments: Not set

Fragment offset: 0
Time to live: 255
Protocol: TCP (6)
Header checksum: 0xcce5 [validation disabled]
[Header checksum status: Unverified]
Source: 10.1.14.1
Destination: 10.1.14.14
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]

Transmission Control Protocol, Src Port: 22008, Dst Port: 23, Seq: 34755, Ack: 93250, Len: 570

R1 - PASSIVE - L3 VPN service - ICMP Type 3/Code 4:

! - as seen from R1 - Passive
! - IP MTU on R2/R3 of 512 bytes is lower than IP packet length and DF bit is set
! - R1 receives ICMP error message from R2
! - note R2 ICMP error message carries Next-Hop MTU
! - "The size in octets of the largest datagram that could be forwarded, along the path of
! the original datagram, without being fragmented at this router. The size includes the
! IP header and IP data, and does not include any lower-level headers."
! - In present L3VPN MPLS-enabled scenario (dual-label) Next-Hop MTU value is 504 bytes
! - In previous MPLS scenario (single-label) Next-Hop MTU value was 508 bytes

2030 0.020299 10.2.3.1 10.1.14.1 ICMP 190 **Destination unreachable
(Fragmentation needed)**

Frame 2030: 190 bytes on wire (1520 bits), 190 bytes captured (1520 bits) on interface 0
Ethernet II, Src: fa:16:3e:5c:f1:80 (fa:16:3e:5c:f1:80), Dst: fa:16:3e:42:18:05
(fa:16:3e:42:18:05)

MultiProtocol Label Switching Header, Label: 24005, Exp: 0, S: 1, TTL: 251

0000 0101 1101 1100 0101 = MPLS Label: 24005
..... 000. = MPLS Experimental Bits: 0
..... 1 = MPLS Bottom Of Label Stack: 1
..... 1111 1011 = MPLS TTL: 251

Internet Protocol Version 4, Src: 10.2.3.1, Dst: 10.1.14.1

```

0100 .... = Version: 4
.... 0101 = Header Length: 20 bytes (5)
Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
Total Length: 172
Identification: 0x002b (43)
Flags: 0x00
  0... .... = Reserved bit: Not set
  .0.. .... = Don't fragment: Not set
  ..0. .... = More fragments: Not set
Fragment offset: 0
Time to live: 253
Protocol: ICMP (1)
Header checksum: 0x9821 [validation disabled]
[Header checksum status: Unverified]
Source: 10.2.3.1
Destination: 10.1.14.1
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Internet Control Message Protocol
  Type: 3 (Destination unreachable)
  Code: 4 (Fragmentation needed)
Checksum: 0xbbac [correct]
[Checksum Status: Good]
Length: 17
[Length of original datagram: 68]
Unused: 0011
MTU of next hop: 504
Internet Protocol Version 4, Src: 10.1.14.1, Dst: 10.1.14.14
  0100 .... = Version: 4
  .... 0101 = Header Length: 20 bytes (5)
  Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
  Total Length: 610
  Identification: 0x7c9f (31903)
  Flags: 0x02 (Don't Fragment)
    0... .... = Reserved bit: Not set
    .1.. .... = Don't fragment: Set
    ..0. .... = More fragments: Not set
  Fragment offset: 0
  Time to live: 255
  Protocol: TCP (6)
  Header checksum: 0xcce5 [validation disabled]
  [Header checksum status: Unverified]
  Source: 10.1.14.1
  Destination: 10.1.14.14
  [Source GeoIP: Unknown]
  [Destination GeoIP: Unknown]
  Transmission Control Protocol, Src Port: 22008, Dst Port: 23, Seq: 586828435, Ack: 754580617

```

PMTUD:TCPオプション(MD5)

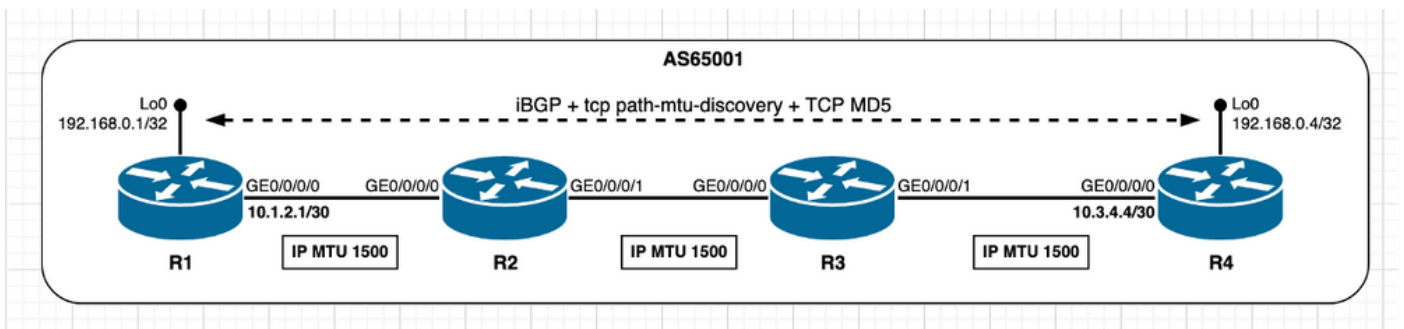


図3.4:PMTUDが有効でTCP MD5認証。

PMTUDの動作については、TCP MD5認証が有効になっている前のシナリオですすでに説明されていたものとは区別されません。Cisco IOS XRでは、以前に使用されているTCP MD5認証と共有したように、追加のオーバーヘッドとアクティブなTCPピアの初期MSS値は同じであると見なします。TCPオプションの使用による影響についての詳細は、前のセクションの「TCPオプションの使用：XRアクティブ」および「TCPオプションの使用：XRパッシブ」を参照してください。このシナリオでのTCP MSSの計算は、次のように要約できます。

- すべてのノードが1500バイトのデフォルトIP MTUを使用します。
- TCPパスMTUディスカバリが有効になっている。
- TCPピアは直接接続されていません。
- TCP MD5認証がR1とR4の両方で有効になっている。R4はBGP接続を管理します。R4は1436バイトのMSSでSYNを送信します。1500 (インターフェイスIP MTU) - 20(minTCP_H) - 20(minIP_H) - 24バイト(IOS XR TCP Options Overhead)。R1はSYNを送信し、MSSが1436バイトのACK。[Received MSS ;ローカルの初期MSS]です。受信MSS 1436バイト。ローカル初期MSS 1460バイト。最小のMSS値が両方のピアで使用されます。

R4から送信されたTCP SYN:

```
! - TCP SYN sourced from R4
```

```
2408  5.695076      192.168.0.4 192.168.0.1 TCP      82      59050  179 [SYN] Seq=0 Win=16384
Len=0  MSS=1436 WS=1
```

```
Frame 2408: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
Ethernet II, Src: fa:16:3e:d7:7e:f6 (fa:16:3e:d7:7e:f6), Dst: fa:16:3e:8f:8f:54
(fa:16:3e:8f:8f:54)
```

```
Internet Protocol Version 4, Src: 192.168.0.4, Dst: 192.168.0.1
```

```
Transmission Control Protocol, Src Port: 59050, Dst Port: 179, Seq: 0, Len: 0
```

```
Source Port: 59050
```

```
Destination Port: 179
```

```
[Stream index: 8]
```

```
[TCP Segment Len: 0]
```

```
Sequence number: 0      (relative sequence number)
```

```
Acknowledgment number: 0
```

```
Header Length: 48 bytes
```

```
Flags: 0x002 (SYN)
```

```
Window size value: 16384
```

```
[Calculated window size: 16384]
```

```
Checksum: 0x20d7 [unverified]
```

```
[Checksum Status: Unverified]
```

```
Urgent pointer: 0
```

```
Options: (28 bytes), Maximum segment size, Window scale, No-Operation (NOP), TCP MD5
```

```
signature, End of Option List (EOL)
```

```
Maximum segment size: 1436 bytes
```

```
Kind: Maximum Segment Size (2)
```

```
Length: 4
```

```
MSS Value: 1436
```

```
Window scale: 0 (multiply by 1)
```

```
No-Operation (NOP)
```

```
TCP MD5 signature
```

```
End of Option List (EOL)
```

TCP SYN, R1からACK:

```
! - TCP SYN,ACK sourced from R1
```

```
2409  0.004352      192.168.0.1 192.168.0.4 TCP      82      179  59050 [SYN, ACK] Seq=0 Ack=1
Win=16384 Len=0  MSS=1436 WS=1
```

```

Frame 2409: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
Ethernet II, Src: fa:16:3e:8f:8f:54 (fa:16:3e:8f:8f:54), Dst: fa:16:3e:d7:7e:f6
(fa:16:3e:d7:7e:f6)
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
Transmission Control Protocol, Src Port: 179, Dst Port: 59050, Seq: 0, Ack: 1, Len: 0
  Source Port: 179
  Destination Port: 59050
  [Stream index: 8]
  [TCP Segment Len: 0]
  Sequence number: 0 (relative sequence number)
  Acknowledgment number: 1 (relative ack number)
  Header Length: 48 bytes
  Flags: 0x012 (SYN, ACK)
  Window size value: 16384
  [Calculated window size: 16384]
  Checksum: 0xcbf8 [unverified]
  [Checksum Status: Unverified]
  Urgent pointer: 0
  Options: (28 bytes), Maximum segment size, Window scale, No-Operation (NOP), TCP MD5
signature, End of Option List (EOL)
    Maximum segment size: 1436 bytes
      Kind: Maximum Segment Size (2)
      Length: 4
      MSS Value: 1436
    Window scale: 0 (multiply by 1)
    No-Operation (NOP)
    TCP MD5 signature
    End of Option List (EOL)

```

R4に表示されるTCPセッションの詳細 – ACTIVE:

! - as seen from R4 - Active

```
RP/0/0/CPU0:R4#show tcp detail pcb 0x121542c0
```

```
Tue Jan 12 13:27:23.526 UTC
```

```
=====
```

```
Connection state is ESTAB, I/O status: 0, socket status: 0
```

```
Established at Tue Jan 12 13:25:41 2021
```

```
PCB 0x121542c0, SO 0x1213c0e4, TCPCB 0x12156010, vrfid 0x60000000,
```

```
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 359
```

```
Local host: 192.168.0.4, Local port: 59050 (Local App PID: 1052958)
```

```
Foreign host: 192.168.0.1, Foreign port: 179
```

```
Current send queue size in bytes: 0 (max 24576)
```

```
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
```

```
Current receive queue size in packets: 0 (max 0)
```

Timer	Starts	Wakeups	Next(msec)
Retrans	6	1	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	3	2	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

```
iss: 3299472269 snduna: 3299473445 sndnxt: 3299473445
```

```
sndmax: 3299473445 sndwnd: 31646 sndcwnd: 4308
```

```
irs: 3225544359 rcvnxt: 3225545535 rcvwnd: 31665 rcvadp: 3225577200
```

SRTT: 89 ms, RTTO: 530 ms, RTV: 441 ms, KRTT: 0 ms
minRTT: 19 ms, maxRTT: 239 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE
Connect retries remaining: 30, connect retry interval: 30 secs

State flags: none
Feature flags: MD5, Win Scale, Nagle, Path MTU
Request flags: Win Scale

Datagrams (in bytes): MSS 1436, peer MSS 1436, min MSS 1436, max MSS 1436

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0

PDU information:
#PDU's in buffer: 0
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:
Num Labels: 0 Label Stack:

RP/0/0/CPU0:R4#

R1に表示されるTCPセッションの詳細 : パッシブ :

! - as seen from R1 - Passive

RP/0/0/CPU0:R1#show tcp detail pcb 0x121560ec
Tue Jan 12 13:25:59.310 UTC

=====
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Tue Jan 12 13:25:31 2021

PCB 0x121560ec, SO 0x121556d4, TCPCB 0x121575bc, vrfid 0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 359
Local host: 192.168.0.1, Local port: 179 (Local App PID: 983326)
Foreign host: 192.168.0.4, Foreign port: 59050

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768) mis-ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer	Starts	Wakeups	Next(msec)
Retrans	3	0	0
SendWnd	0	0	0
TimeWait	0	0	0
AckHold	3	2	0
KeepAlive	1	0	0
PmtuAger	0	0	0
GiveUp	0	0	0
Throttle	0	0	0

iss: 3225544359 snduna: 3225545516 sndnxt: 3225545516
sndmax: 3225545516 sndwnd: 31684 sndcwnd: 4308

```
irs: 3299472269 rcvnext: 3299473426 rcvwnd: 31665 rcvadv: 3299505091
```

```
SRTT: 37 ms, RTTO: 300 ms, RTV: 244 ms, KRTT: 0 ms  
minRTT: 9 ms, maxRTT: 239 ms
```

```
ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30 sec  
Giveup time: 0 ms, Retransmission retries: 0, Retransmit forever: FALSE  
Connect retries remaining: 0, connect retry interval: 0 secs
```

```
State flags: none  
Feature flags: MD5, Win Scale, Nagle, Path MTU  
Request flags: Win Scale
```

```
Datagrams (in bytes): MSS 1436, peer MSS 1436, min MSS 1460, max MSS 1460
```

```
Window scales: rcv 0, snd 0, request rcv 0, request snd 0  
Timestamp option: recent 0, recent age 0, last ACK sent 0  
Sack blocks {start, end}: none  
Sack holes {start, end, dups, rxmit}: none
```

```
Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO  
Socket states: SS_ISCONNECTED, SS_PRIV  
Socket receive buffer states: SB_DEL_WAKEUP  
Socket send buffer states: SB_DEL_WAKEUP  
Socket receive buffer: Low/High watermark 1/32768  
Socket send buffer : Low/High watermark 2048/24576, Notify threshold 0
```

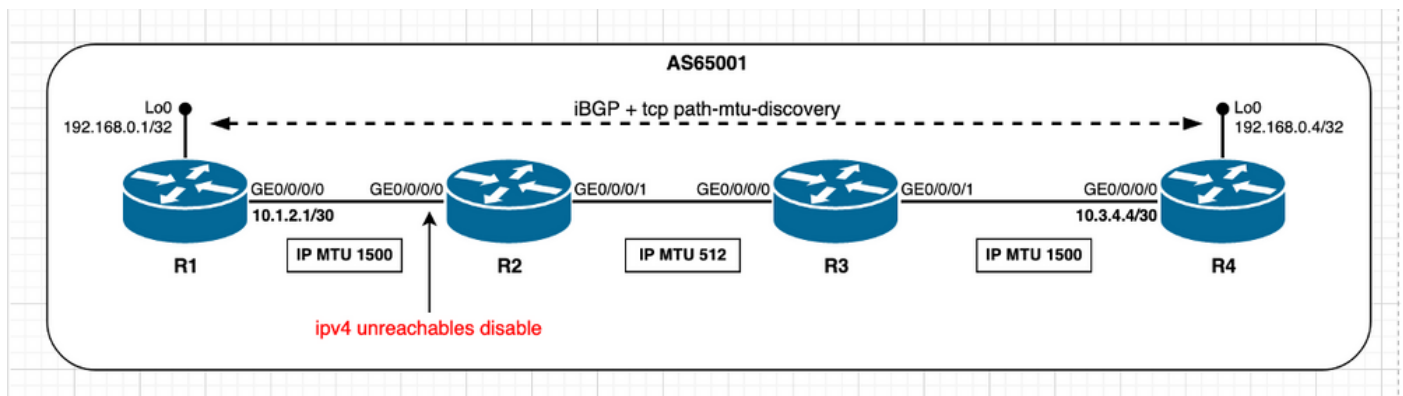
```
PDU information:  
#PDU's in buffer: 0  
FIB Lookup Cache: IFH: 0x40 PD ctx: size: 0 data:  
Num Labels: 0 Label Stack:
```

```
RP/0/0/CPU0:R1#
```

PMTUD – ブラックホール検出

「PMTUD – パスセグメントのIP MTUが小さい」の項で前述したように、有効な場合のTCP PMTUDはICMP (宛先到達不能 – タイプ3) の受信によってトリガーされます。Fragmentation needed - Code 4)メッセージ。これらのメッセージが何らかの理由で受信されず、PMTUDがトリガーされない場合があります。この場合、TCPピア間のパスの最小IP MTUは学習されません。このようなシナリオでは、IPパケットにDFビットが設定されていて、最小のIP MTUパスセグメントよりも大きいサイズが設定されている場合、潜在的なブラックホールが発生します。これらのパケットはサイレントドロップされます。

このセクションでは、Cisco IOS XRがこのような潜在的なブラックホールのシナリオを検出して動作する方法を説明します。この目的のために、次の図とCLIの出力に示すように、R2インターフェイスGE0/0/0/0ではIPv4到達不能の機能が無効になっています。



イメージ3.5:R1/R4およびR2 IPv4到達不能でPMTUDが有効になっている。

R2で無効にされたIPv4到達不能：

```
!- R2 - IP unreachable is disabled
```

```
RP/0/0/CPU0:R2#show run interface gigabitEthernet 0/0/0/0
```

```
Thu May 13 12:09:45.483 UTC
```

```
interface GigabitEthernet0/0/0/0
```

```
ipv4 address 10.1.2.2 255.255.255.252
```

```
ipv4 unreachable disable
```

```
!
```

```
RP/0/0/CPU0:R2#show ipv4 interface gigabitEthernet 0/0/0/0
```

```
Thu May 13 12:10:04.112 UTC
```

```
GigabitEthernet0/0/0/0 is Up, ipv4 protocol is Up
```

```
Vrf is default (vrfid 0x60000000)
```

```
Internet address is 10.1.2.2/30
```

```
MTU is 1514 (1500 is available to IP)
```

```
Helper address is not set
```

```
Multicast reserved groups joined: 224.0.0.2 224.0.0.1 224.0.0.5
```

```
224.0.0.6
```

```
Directed broadcast forwarding is disabled
```

```
Outgoing access list is not set
```

```
Inbound common access list is not set, access list is not set
```

```
Proxy ARP is disabled
```

```
ICMP redirects are never sent
```

```
ICMP unreachable are never sent
```

```
ICMP mask replies are never sent
```

```
Table Id is 0xe0000000
```

このブラックホールのシナリオに対するCisco IOS XRの対処方法は、同じパケットを2回再送信し、それでも失敗した場合は、期待されるTCP ACKを受信せずに再試行し、パスMTUディスカバリーで次に明確な値を使用します(PlateusのリストのIP MTUが低い)。要約すると、Cisco IOS XRでは、サイズが大きいためパケットが宛先までのパスのどこかでドロップされ、パケットの再送信を介してパケットを回避しようとする場合が考えられます。この動作は、次の例でノードR1インターフェイスで取得したパケットキャプチャとdebug tcp pmtudの出力から確認できます。

R1でのIOS-XRブラックホール検出：

```
! - at R1
```

```
! - Original BGP Update message is sent
```

```
! - Note IP Total Length of 1116 bytes and TCP Segment Length of 1076 bytes
```

```
! - R2 filters such packet and send and ICMP error message towards R1 which triggers PMTUD
```

```
! - But because IPv4 unreachable are disabled at R2 GE0/0/0/0 ICMP message is not sent
```

```
! - Hence BGP message is silently filtered at R2
```

```
562      7.638774      192.168.0.1 192.168.0.4 BGP      1130      UPDATE Message, KEEPALIVE Message
```

```
Frame 562: 1130 bytes on wire (9040 bits), 1130 bytes captured (9040 bits) on interface 0
```

```
Ethernet II, Src: fa:16:3e:42:18:05 (fa:16:3e:42:18:05), Dst: fa:16:3e:5c:f1:80
```

```
(fa:16:3e:5c:f1:80)
```

```
Internet Protocol Version 4, Src: 192.168.0.1, Dst: 192.168.0.4
```

```
0100 .... = Version: 4
```

```
.... 0101 = Header Length: 20 bytes (5)
```

```
Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
```

```
Total Length: 1116
```

```
Identification: 0x4a37 (18999)
```

```
Flags: 0x02 (Don't Fragment)
```

```
0... .... = Reserved bit: Not set
```

```

    .1.. .... = Don't fragment: Set
    ..0. .... = More fragments: Not set
Fragment offset: 0
Time to live: 255
Protocol: TCP (6)
Header checksum: 0x229b [validation disabled]
[Header checksum status: Unverified]
Source: 192.168.0.1
Destination: 192.168.0.4
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
Transmission Control Protocol, Src Port: 179, Dst Port: 57082, Seq: 318, Ack: 251, Len: 1076
Border Gateway Protocol - UPDATE Message
Border Gateway Protocol - KEEPALIVE Message
<snip>

! - at R1
! - No TCP ACK is received
! - Packet retransmission is attempted (2 attempts)
! - Note initial MSS value is of 1460 bytes

563    0.560058      192.168.0.1 192.168.0.4 TCP      1130    [TCP Retransmission] 179  57082
[PSH, ACK] Seq=318 Ack=251 Win=32593 Len=1076
564    1.101367      192.168.0.1 192.168.0.4 TCP      1130    [TCP Retransmission] 179  57082
[PSH, ACK] Seq=318 Ack=251 Win=32593 Len=1076

! - at R1
! - Still no TCP ACK received; previous retransmissions failed
! - Next lower plateau value is attempted - 1492 bytes
! - Packet retransmission is attempted (2 attempts)

RP/0/0/CPU0:May 13 10:20:44.251 UTC: tcp[399]: [t1] PCB 0x15392224: Trying next lower MTU: 1452

567    1.850294      192.168.0.1 192.168.0.4 TCP      1130    [TCP Retransmission] 179  57082
[PSH, ACK] Seq=318 Ack=251 Win=32593 Len=1076
568    1.111361      192.168.0.1 192.168.0.4 TCP      1130    [TCP Retransmission] 179  57082
[PSH, ACK] Seq=318 Ack=251 Win=32593 Len=1076

! - at R1
! - Still no TCP ACK received; previous retransmissions failed
! - Next lower plateau value is attempted - 1006 bytes
! - Packet retransmission is attempted (2 attempts)

RP/0/0/CPU0:May 13 10:20:47.560 UTC: tcp[399]: [t1] PCB 0x15392224: Trying next lower MTU: 966

569    2.198327      192.168.0.1 192.168.0.4 TCP      1020    [TCP Retransmission] 179  57082
[ACK] Seq=318 Ack=251 Win=32593 Len=966
570    1.109602      192.168.0.1 192.168.0.4 TCP      1020    [TCP Retransmission] 179  57082
[ACK] Seq=318 Ack=251 Win=32593 Len=966

! - at R1
! - Still no TCP ACK received; previous retransmissions failed
! - Next lower plateau value is attempted - 508 bytes
! - Original information (TCP Length of 1076 bytes) is split in three distinct packets
! - TCP Segment Lengths 468 + 468 + 140 = 1076
! - TCP ACK is received from peer R4

RP/0/0/CPU0:May 13 10:20:50.870 UTC: tcp[399]: [t1] PCB 0x15392224: Trying next lower MTU: 468

571    2.205552      192.168.0.1 192.168.0.4 TCP      522     [TCP Retransmission] 179  57082
[ACK] Seq=318 Ack=251 Win=32593 Len=468
573    0.004254      192.168.0.1 192.168.0.4 TCP      522     [TCP Retransmission] 179  57082
[ACK] Seq=786 Ack=251 Win=32593 Len=468
574    0.002724      192.168.0.1 192.168.0.4 TCP      194     [TCP Retransmission] 179  57082

```


[PSH, ACK] Seq=1254 Ack=251 Win=32593 **Len=140**

! - Peer R4 TCP ACK is received

575 0.223172 192.168.0.4 192.168.0.1 TCP 54 57082 179 [ACK] Seq=251 Ack=1394
Win=31469 Len=0