応募区分：研究型論文

IOS Embedded Packet Capture(EPC) 機能検証結果報告

高原　也寿明（たかはら　やすあき）
伊藤忠テクノソリューションズ株式会社
ネットワークインフラ技術推進部

■ 要約

シスコ社ブランチ・ルーター ISR の IOS 12.4(20)T から、Embedded Packet Capture (EPC) と呼ばれる機能が実装された。
これは、ルーターが取り扱う IPv4/IPv6 パケットをキャプチャする機能であり、キャプチャしたパケットは、CLI コマンドで表示させたり、PCAP ファイルとして出力させ、Wireshark 等のパケット解析アプリケーションで取り扱うことができる。

障害解析時に非常に有用なツールではあるが、細かい仕様や動作を説明したドキュメントが少ないため、客先等の現場ではなかなか使用されないのが実情である。

そこで検証によって IOS EPC の仕様を明らかにし、現場でこの機能が使用できるよう、マニュアルを作成した。マニュアルは 40 ページを超えるためここには掲載せず、作成の過程で判明した動作及び仕様を報告する。

目次

1. 背景

シスコ社ブランチ・ルーター ISR の IOS 12.4(20)T から、Embedded Packet Capture (EPC) と呼ばれる機能が実装された。これは、ルーターが取り扱う IPv4/IPv6 パケットをキャプチャする機能であり、キャプチャしたパケットは、CLI コマンドで表示させたり、.pcap ファイルとして出力させ、Wireshark 等のパケット解析アプリケーションで取り扱うことができる。

障害解析のツールとしてこの機能を使用する場合、以下のような利点がある。

・ルーターの接続環境を変更する必要が無い。
例えば、ケーブルで直結したルーター間を流れるパケットをキャプチャしようとする時、従来であれば必要なケーブルを一度外し、スイッチを間に挟んで SPAN 設定を行い、測定装置を接続するといった作業が不要になる。このことは、サービスの停止を伴う作業を排除すると共に、障害解析作業に速やかに入れることを意味する。

・イーサーネット以外のメディアを流れるパケットがキャプチャできる。
ブランチ・ルーターは、シリアルインターフェイス等、イーサーネット以外のメディアもサポートしている。このようなメディアを流れるパケットを解析したい場合、従来であればラインモニター等、特殊な（そして高価な）装置を間にはさんでパケットをキャプチャする必要があった。EPC によって、このような束縛からも解放される。

しかしながら、実際に客先等、緊急度の高い現場においてこの機能を使おうとした場合、以下のようなことが妨げとなる。

・CPU やメモリーへの負荷が不安。
・とりこぼしや、送受信順序の狂いなどが生じないかが不安。
・ルーター自身宛、又はルーター自身が発生させるパケットの扱いが不明。
・アクセスリストが適用されていた場合のキャプチャ動作が不明。
・QoS を適用し、帯域を絞っている場合のキャプチャ動作が不明。
・NAT を行っている場合のキャプチャ動作が不明。
・GRE のトンネルを使用した場合のキャプチャ動作が不明。
・そもそも使用方法がよくわからない。

つまり、現場で使用するには、不明要素や不安要素が多いのである。

そこで、Cisco 2911、IOS 15.1.4M8 を使用して EPC 機能を検証し、マニュアルを作成した。
ここでは、検証時に判明した興味深い事項及び考慮事項について記述する。

2. キャプチャ条件

2.1. キャプチャできるパケット

EPC では、ルーターが受信、又は送出する IPv4 及び IPv6 パケットのキャプチャが可能である。ユニキャストのみならず、ブロードキャスト及びマルチキャストパケットもキャプチャ可能であることを確認した。ただし、ルーターが転送するパケットをキャプチャするには、CEF が有効になっている必要がある。転送する IPv6 パケットをキャプチャするには、もちろん IPv6 CEF が有効になっていなければならない。
※IPv6 CEF を有効にするには、IPv4 CEF が有効になっている必要がある。

ルーター管理用パケットや、ルーティングプロトコル等、ルーター自身に宛てたパケットやルーター自身が生成し送出するパケットもキャプチャ可能であることを確認した。これらのパケットについては、CEF の有無にかかわらずキャプチャが可能であった。

2.2. キャプチャできないパケット

ARP, CDP 等、IPv4/IPv6 ではないパケットはキャプチャできない。また、PPPoE 等、レイヤ 2 で示す上位プロトコルが IP で無いものについては、たとえ IP パケットを運んでいてもキャプチャされないことを確認した。

2.3. 対象インターフェイス

基本的には CEF で IP 転送を行うインターフェイスで受信、送出したパケットをキャプチャ対象とすることができる。物理インターフェイスはもちろん、802.1Q のサブインターフェイス、GRE 等のトンネルインターフェイス、PPPoE 等で使用する Dialer インターフェイスを流れるパケットもキャプチャ可能である。

802.1Q のサブインターフェイスの場合、TAG がついた状態でパケットをキャプチャすることができる。トンネルインターフェイスの場合、受信側ではトンネルヘッダが外された状態でパケットをキャプチャし、送出側ではトンネルヘッダを付加した状態でキャプチャする。

Loopback インターフェイスはキャプチャの対象にはならない。Loopback インターフェイス宛、又は Loopback インターフェイス発のパケットキャプチャはインターフェイスを指定せず、Process Switching で処理されたパケットをキャプチャし、後から IP アドレス等で絞り込む必要がある。

特殊なインターフェイスとして、"drop" が指定可能。これは、CEF 処理により破棄されたパケットをキャプチャするためのもの。経路の不在による破棄、ACL による破棄、Null0 宛てパケットの破棄などが該当する。輻輳等により、送出インターフェイスの Queue で破棄されたパケットや、ICMP の生成トリガーになったパケットは CEF の取り扱い対象外となるため "drop" インターフ

ェイスのキャプチャ対象にはならない。

3. 指定取得タイプと取得動作

3.1. CEF In 指定時の取得動作

この取得タイプでは、指定したインターフェイスで受信し、CEF で取り扱うパケットをキャプチャする。基本的にはルーターが転送するユーザートラフィックのキャプチャとなる。

このキャプチャは、他の機能よりも先に行われる。従って、Ingress ACL によって破棄の対象となっているパケットや、経路が存在しないために破棄されるパケット、Ingress Policer によって破棄されるパケットもキャプチャされる。

また、EPC ではキャプチャしたパケットの入力及び出力インターフェイスを記録するが、このタイプで取得したパケットは転送処理の前にキャプチャされるため、入力インターフェイスは記録されるが、出力インターフェイス情報は記録されず、必ず "None" となる。

QoS によって入力パケット DSCP の書き換え等が指示されていた場合や NAT による IP Address 変更が行われる場合も、キャプチャが QoS や NAT の処理前に行われるため、変更前（受信時オリジナル）の状態でキャプチャされる。

なお、ルーター自身を宛先としたトラフィックの一部もキャプチャ可能である。

外部から機器を管理するためのプロトコル、Telnet, SSH, Ping(Unicast/Multicast/Broadcast), SNMP, NTP がキャプチャされることを確認した。

同じルーター自身宛であっても、ルーティングプロトコル、RIP, EIGRP, OSPF, BGP, PIM については、この取得タイプではキャプチャできないことを確認した。※

※RIP, BGP については、最初の 1 パケットのみキャプチャされた。これは UDP 又は TCP を使用するトラフィックについては、ルーティングプロトコルのセッションであることを確認できるまでは他のルーター宛パケットと同じ扱いになるためと推測される。

3.2. CEF Out 指定時の取得動作

この取得タイプでは、指定したインターフェイスに、CEF によって受信インターフェイスから転送され、送出されるパケットをキャプチャする。ルーターが転送するユーザートラフィックのキャプチャとなる。ルーター自身が生成して送出するパケットはこの指定ではキャプチャすることはできない。

このキャプチャは、他の処理が終了し、インターフェイスの送出 Queue に送り込まれる前に行われる。このため、Egress ACL や Egress Policer によって破棄されたパケットはキャプチャされない。ただし、輻輳による送出 Queue 溢れや、Shaping による帯域制限によるパケット破棄はキャプチャ後に行われる。

NAT や QoS によるパケット内容の変更がある場合、変更後のパケットがキャプチャされる。

この指定でキャプチャされたパケットの入出力インターフェイス情報は、送受信インターフェイスが両方とも記録される。

3.3. Process-Switched In 指定時の取得動作

この取得タイプでは、Process Switching 処理で受信したパケットをキャプチャする。ルーターの通常プロセス（割り込みではないプロセス）で処理するパケットが対象となり、通常はルーティングプロトコルやルーター宛管理トラフィック等、ルーター自身を宛先にしたパケットがキャプチャされる。

ただし、何らかの理由で CEF が無効になっていた場合は、ルーターが転送するユーザートラフィックもキャプチャされる。

Ingress Control Plane Policing (CoPP) 処理が行われた後にキャプチャされるため、Ingress CoPP で破棄されたパケットはキャプチャできない。

この取得タイプではインターフェイスの指定はできず、どのインターフェイスから入ったパケットであってもキャプチャされる。入力インターフェイス名は記録されているため、CLI でキャプチャ内容を直接表示させる場合は入力インターフェイス名でフィルターをかけることができる。

興味深い動作として、発信元ホストに ICMP による通知を行う場合、その原因となったパケットは Process Switching 処理になるためキャプチャ対象となるが、単に破棄されたパケットは CEF による処理となって、キャプチャ対象とはならないことが挙げられる。

例えば、キャプチャを行う機器に ACL を適用し、破棄の対象とした IOS 機器からキャプチャを行っている機器に Ping を行い、以下の結果を得たとすると、5 つのパケットの内、キャプチャ機器が ICMP Unreachable を生成したために結果が "U" と表示された 3 つの ICMP Echo がキャプチャされていることになる。

※IOS のディフォルトでは、1 インターフェイスからの ICMP 送出を 500 msec 間隔に制限しているため、ICMP を一度生成すると、500 msec 経過するまでは ICMP 生成を行わない。

```
Router#ping 172.30.30.252
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.30.30.252, timeout is 2 seconds:
U.U.U
```

3.4. Process-Switched Out 指定時の取得動作

この取得タイプでは、Process Switching 処理で送出したパケットをキャプチャする。ルーターの通常プロセス（割り込みではないプロセス）で処理するパケットが対象となり、通常はルーティングプロトコルやルーター宛管理トラフィック等、ルーター自身が生成したパケットがキャプチャされる。

ただし、何らかの理由で CEF が無効になっていた場合は、ルーターが転送するユーザートラフィックもキャプチャされる。

この取得タイプではインターフェイスの指定はできず、どのインターフェイスに出力するパケットであってもキャプチャされる。送出インターフェイス名は記録されているため、CLI でキャプチャ内容を直接表示させる場合は送出インターフェイス名でフィルターをかけることができる。

キャプチャは、Egress CoPP 処理前に行われるため、Egress CoPP で破棄されるパケットもキャプチャされる。

4. ポイントとバッファ

4.1. ポイントとは

EPC における「ポイント」とは、キャプチャ動作を操作するための「操作ポイント」である。パケットの取得タイプや、どのバッファにキャプチャ結果を流し込むかを指定しておいて、任意のタイミングで開始／停止させることができる。

1 ポイントには任意の 1 インターフェイスか、全インターフェイスのみが指定可能だが、複数のポイントを作成することで、任意の複数インターフェイスでのパケットキャプチャが可能になる。

4.2. バッファとは

EPC における「バッファ」とは、キャプチャしたパケットを収容するためにメインメモリー内に確保される領域である。ポイントの指示によってキャプチャされたパケット情報は、このバッファに格納される。

複数のポイントがキャプチャした結果を 1 バッファで収容することもできるし、複数のバッファを作成し、ポイント毎に結果を分けて収容することもできる。

5. バッファ使用形態

5.1. バッファサイズ

キャプチャパケットを収容するバッファは、256 ～ 102,400 (Kbyte) の範囲で指定でき、メインメモリー内に確保される。ディフォルトは 1,024 (Kbyte) である。

5.2. パケット収容範囲

パケットは、レイヤ 2 のヘッダがついた状態でバッファに収容される。最後の FCS 部分は収容されない。例えば、イーサーネットの TAG 無し IP パケットを収容する場合、IP 部分に 14-Bytes のイーサーネットヘッダーが加えられた状態で収容される。

5.3. 収容パケット長

ディフォルトでは、パケットの先頭から 68-Bytes までが収容対象となっている。これを超えた部分はキャプチャされないため、バッファに取り込めるパケット数が増加する。ヘッダ部分のみで解析を行うような場合に有効であると思われる。

パケットの足きりを避けるのであれば、最大 9,500-Bytes まで収容対象部を拡張することができる。

5.4. パケット収容数の目安

バッファには、タイムスタンプや送受信インターフェイス名等の付帯情報が加えられてキャプチャパケットが収容される。固定長のパケットをキャプチャさせた場合に保持できたパケット数を記録し、パケット長やバッファサイズを変更して調査を行った結果、概ね以下のことが判明した。

・バッファは 8-Byte 単位で消費される。(例 : 9-Byte 収容時、16-Byte 必要)
・各パケットには 64-Byte の付帯情報が付加される。

以下表 1 に、パケット長が固定であり、収容パケット長設定による足きりをしない状態で、ディフォルトの 1,024-Kbyte バッファに収容できるパケット数を示す。

表 1

1024 Kbytes バッファに収容できるパケット数	
IP 46-Byte (Ether 60-Byte)	8,192 個
IP 100-Byte (Ether 114-Byte)	5,698 個
IP 200-Byte (Ether 214-Byte)	3,744 個
IP 500-Byte (Ether 514-Byte)	1,795 個
IP 1,000-Byte (Ether 1014-Byte)	970 個

※EPC は、バッファの空き容量が少なくなり、キャプチャしたパケットを完全に収納できない場合でも、そのパケットを頭から可能な分だけ収容し、入りきらない部分を破棄する。表 1 は、そのようにして足きりされた最後のパケットをカウントから除いている。

6. バッファ内容の表示

6.1. タイムスタンプ

キャプチャされたパケットのタイムスタンプは、4-msec 毎にしか記録されない。これは、キャプチャしたタイミングでリアルにタイムスタンプを付加するのではなく、キャプチャパケットを 4 msec 毎にバッファに書き込み、その時にタイムスタンプを付加するためと思われる。このため、実際にキャプチャされた時刻と、バッファに書き込まれたタイムスタンプの間には誤差が生じる。

なお、タイムスタンプで示される時刻が同じであっても、その 4-msec の間にキャプチャされたパケット間に順序の狂いは全く確認されなかった。タイムスタンプの粒度が荒いだけで、バッファに格納されたパケット順序が狂うことはないと考えられる。

以下が、タイムスタンプを含むバッファ内容の表示例である。

```
07:14:48.571 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.571 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.571 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.571 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.571 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.575 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.579 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.579 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
07:14:48.579 UTC Jul 3 2014 : IPv4 LES CEF     : Gi0/1 None
```

6.2. ダンプ表示の考慮点

バッファに収納されたパケット内容は HEX ダンプ形式で表示させることができる。以下が表示例である。

```
03:55:47.153 UTC Jul 3 2014 : IPv4 LES CEF    : Gi0/1 None

30DDD8D0:           5475D0F7 8A115475 D0F76031      TuPw..TuPw`1
30DDD8E0: 08004500 00640044 0000FE01 3925AC14  ..E..d.D..~.9%,.
30DDD8F0: 0C01AC1E 1EFC0800 55F30014 00000000  ..,...|..Us.....
30DDD900: 00000EDB 1968ABCD ABCDABCD ABCDABCD  ...[.h+M+M+M+M+M
30DDD910: ABCDABCD ABCDABCD ABCDABCD ABCDABCD  +M+M+M+M+M+M+M+M
30DDD920: ABCDABCD ABCDABCD ABCDABCD ABCDABCD  +M+M+M+M+M+M+M+M
30DDD930: ABCDABCD ABCDABCD ABCDABCD ABCDABCD  +M+M+M+M+M+M+M+M
30DDD940: ABCDABCD ABCD00                      +M+M+M.
```

表示は 4 Bytes 毎に区切られて 1 行あたり 16-Byte で表示される。ただし、パケットの最初が必ず一番左の 0 番地から始まるとは限らない。0,4,8,C 番地のいずれかから開始される。上記の例では 2 ブロック目の 30DDD8D4 番地から表示されている。

一番左のインデックス番号はバッファ内に格納されているメモリー番地等とは無関係であり、表示時にダンプデータの位置がわかりやすいよう、一時的に展開された番号にすぎない。以下は同じパケットを再度表示させた例である。まったく違うインデックス番号になっている。

```
03:55:47.150 UTC Jul 3 2014 : IPv4 LES CEF    : Gi0/1 None

2B4BC0D0:                 5475D0F7 8A115475        TuPw..Tu
2B4BC0E0: D0F76031 08004500 00640044 0000FE01  Pw`1..E..d.D..~.
2B4BC0F0: 3925AC14 0C01AC1E 1EFC0800 55F30014  9%,...,..|..Us..
2B4BC100: 00000000 00000EDB 1968ABCD ABCDABCD  .......[.h+M+M+M
2B4BC110: ABCDABCD ABCDABCD ABCDABCD ABCDABCD  +M+M+M+M+M+M+M+M
2B4BC120: ABCDABCD ABCDABCD ABCDABCD ABCDABCD  +M+M+M+M+M+M+M+M
2B4BC130: ABCDABCD ABCDABCD ABCDABCD ABCDABCD  +M+M+M+M+M+M+M+M
2B4BC140: ABCDABCD ABCDABCD ABCD00             +M+M+M+M+M.
```

ダンプデータの最後の 1-Byte "00" は「ゴミ」であり、実際のパケット内容では無い。現在のところ、CLI によるパケットダンプ表示時には必ず付加されてしまうので、このダンプ表示を使用して IP データ部分の解析を行う場合には注意が必要である。

7. パフォーマンスへの影響

ブランチ・ルーター ISR は、ソフトウェア処理でパケット転送を行っているため、CPU の使用率が上がればパフォーマンスに影響が発生する可能性がある。そこで、EPC が CPU 及びパフォーマンスに与える影響についても検証を行った。

検証では、Cisco 2911 本体にある 3 インターフェイスの内、2 ポートを使用して固定長パケットを転送させ、パフォーマンスと CPU 使用率の関係をチェックした。また、残りの 1 ポートを使用して、別ルーターと OSPF の隣接関係を確立させ、CPU 高負荷時に他の処理に与える影響もチェックした。なお、ルーターに転送トラフィックを流さない時の CPU 使用率はほとんど 0％ であった。

まず、EPC を使用しない状態でのパフォーマンス計測を行った。この時、"show processes cpu sorted" コマンドによる CPU 負荷状況のチェックを行い、パケット転送による負荷が全て割り込み処理で効率よく行われていることを確認した。

割り込み処理による CPU 使用率が概ね 90% を超えると、受信パケットの一部が "Overrun Error" となって受信処理できなくなり、パフォーマンスの限界に達することが確認できた。また、このとき OSPF の隣接関係が失われる事象が発生することがあることも観察でき、ルーターが行っている他の処理にも影響を及ぼす可能性があることが確認できた。

また、CPU 負荷によるパフォーマンスへの影響は、トラフィック量とパケットサイズに左右されることが確認できた。

EPC を有効にした時の CPU の使用率とパフォーマンスへの影響を確認した。この時上昇した CPU の使用形態が全て割り込み処理であることを確認した。興味深いことに、ポイントを作成しておくだけで、実際にキャプチャを行わない場合でもある程度 CPU 使用率が上昇することが判明した。

次の表 2 はこの検証で確認された EPC の有無によるパフォーマンスの限界値（Mbps）とパケットサイズの関係である。

表 2

パフォーマンス上限			
パケット長	EPC 無し	EPC 非アクティブ	EPC アクティブ
64-Bytes	172 Mbps	148 Mbps	99 Mbps
128-Bytes	345 Mbps	291 Mbps	189 Mbps
256-Bytes	690 Mbps	584 Mbps	384 Mbps
512-Bytes	over 1Gbps	over 1Gbps	575 Mbps
768-Bytes	over 1Gbps	over 1Gbps	728 Mbps
1024-Bytes	over 1Gbps	over 1Gbps	816 Mbps
1518-Bytes	over 1Gbps	over 1Gbps	over 1Gbps

以下の図 1, 2, 3 は、CPU 使用率とトラフィック量の計測結果の一部をグラフ化したものである。

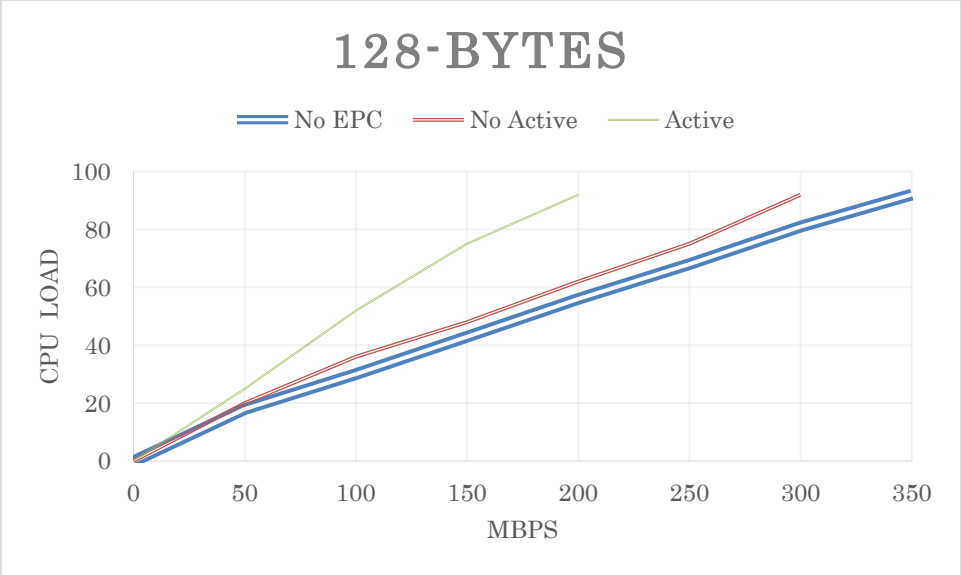

図 1

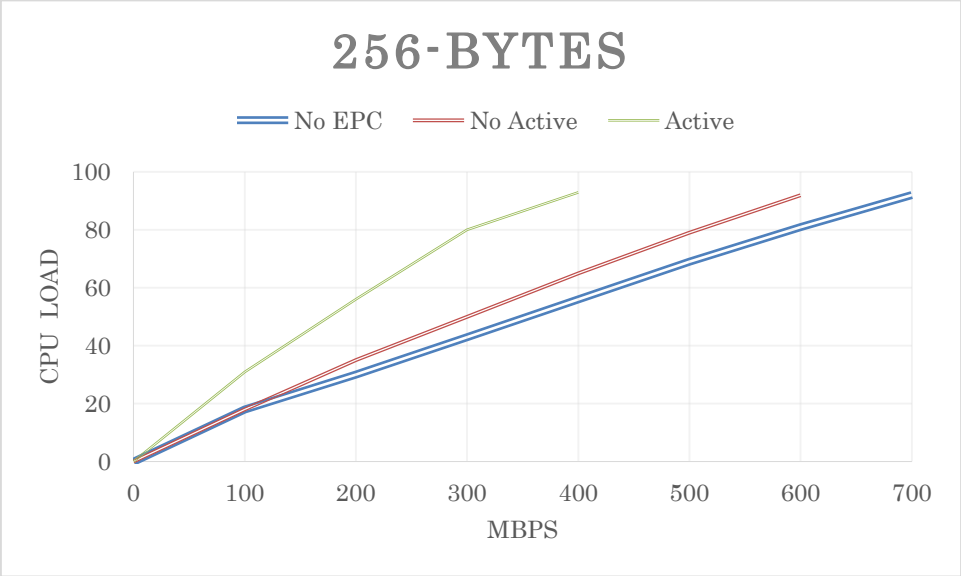

図 2

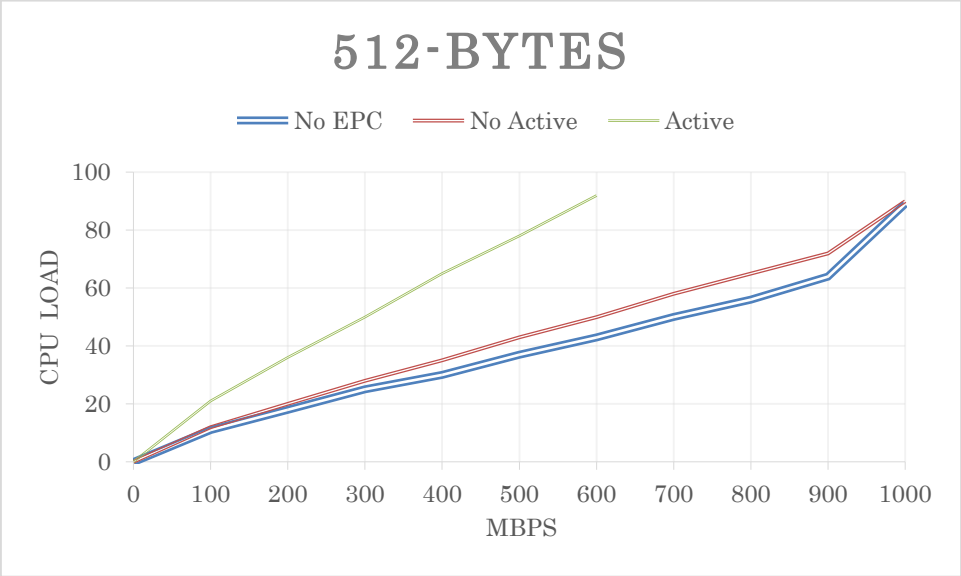

図 3

8. PCAP ファイル

バッファに収容された内容は、機器自身の Flash メモリーや TFTP サーバー等に PCAP ファイルとして出力することができ、Wireshark 等の解析アプリケーションで取り扱うことができるようになる。しかしながら、PCAP ファイルにすると、レイヤ 2 の情報が失われてしまい、IP 部分のみしか解析できないようになってしまうことを確認した。

この動作について、ルーターはイーサーネット以外のメディアも取り扱えるが、PCAP ファイルを扱うアプリケーションの大部分はおそらくイーサーネットのみをサポートしており、サポートしていないレイヤ 2 の情報が読み込まれることで、予期せぬ不具合が発生することを避けるための仕様ではないかと推測している。

9. 収容制限

ポイントとバッファの数には以下の制限があることを確認した。

・バッファ最大数 10
・1 バッファを共有できるポイント最大数 6
・1 インターフェイスが所属できるポイント最大数 8

10. 所感

基本的な動作と使用方法を知っていれば、EPC によるパケットキャプチャは、トラブルシューティングや動作の確認に非常に有用なツールであると感じた。今後も状況に応じて積極的に使用していきたいと考えている。場合によっては EEM と組み合わせ、あるトリガーによってパケットキャプチャを開始するような運用方法も考えられると思う。

ただし、高い CPU 使用率で運用されている機器で実行することはリスクが伴うことを知っておくべきであろう。また、EPC のポイントを残しておくと機器の負荷が高いままになってしまうため、EPC の使用が終了したら、ポイント停止だけではなく消去も確実に行うことを徹底したい。

いまの実装でも充分満足しているが、今後、もしできることであれば、Catalyst4500 に搭載された Wireshark のように、CLI で直接パケットの解析表示ができるようになると嬉しい。

また、PCAP ファイルに出力する際、せめてイーサーネットメディアのキャプチャ結果だけでも、レイヤ 2 の情報が残せるようになるとさらに良いと感じた。

11. 参照資料

［1］Embedded Packet Capture Configuration Guide, Cisco IOS Release 15M&T
http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/epc/configuration/15-mt/epc-15-mt-book/nm-packet-capture.html
［2］Cisco IOS Embedded Packet Capture Command Reference
http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/epc/command/epc-cr-book.html