Elevato utilizzo della CPU sugli switch Catalyst a causa del traffico multicast IPv6

Sommario

Introduzione

In questo documento viene descritto l'elevato utilizzo della CPU su diverse piattaforme Catalyst a causa del sovraccarico dei pacchetti di rilevamento del listener multicast IPV6 e vengono illustrati i metodi per risolvere il problema.

Prerequisiti

Non sono previsti prerequisiti.

Requisiti

Nessun requisito specifico previsto per questo documento.

Componenti usati

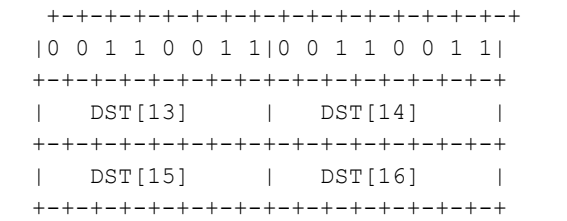
Per la stesura del documento, sono stati usati switch Cisco Catalyst serie 6500, Catalyst serie 4500 e Catalyst serie 3850.

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti.

Problema

L'utilizzo elevato della CPU può essere rilevato su alcune piattaforme Cisco Catalyst a causa del traffico multicast IPv6 con indirizzo MAC compreso nell'intervallo 333.xxxx.xxxx puntato alla CPU.

In base a RFC7042, tutti gli identificatori multicast MAC-48 con prefisso "33-33" (ovvero, i 2**32 identificatori MAC multicast compresi nell'intervallo da 33-33-00-00-00-00 a 33-33-FF-FF-FF) vengono utilizzati come specificato in [RFC2464] per il multicast IPv6. Un pacchetto IPv6 con un indirizzo di destinazione multicast DST, costituito dai sedici ottetti da DST[1] a DST[16], viene trasmesso all'indirizzo multicast Ethernet i cui primi due ottetti corrispondono al valore 3333 esadecimale e gli ultimi quattro ottetti corrispondono agli ultimi quattro ottetti di DST, come mostrato nella Figura 1.



(Figure 1: Address Mapping - Multicast)

In alcune occasioni è stato rilevato che quando i dispositivi host che utilizzano una determinata scheda NIC passano alla modalità di sospensione, inondano il traffico multicast IPv6. Questo problema non è limitato a un particolare fornitore di host, sebbene alcuni chipset mostrino questo comportamento più spesso di altri.

Risoluzione dei problemi e soluzione

Per verificare se lo switch Catalyst che rileva un utilizzo elevato della CPU è interessato dal problema e implementare le rispettive soluzioni, è possibile utilizzare le procedure seguenti.

Switch Catalyst serie 3850

Sugli switch Catalyst 3850, NGWC L2M Process utilizza la CPU per elaborare i pacchetti IPv6. Quando lo snooping MLD (Multicast Listener Discovery) è disabilitato sullo switch, il pacchetto MLD join/leave viene inviato a tutte le porte membro. Inoltre, se sono presenti molti pacchetti MLD in arrivo con join/leave, questo processo richiederà più cicli della CPU per l'invio dei pacchetti su tutte le porte membro. È stato rilevato che quando alcuni computer host passano alla modalità di sospensione, possono inviare diverse migliaia di pacchetti al secondo di traffico MLD IGMPv6.

```
3850#show processes cpu detailed process iosd sorted | exc 0.0
Core 0: CPU utilization for five seconds: 43%; one minute: 35%; five minutes: 33%
Core 1: CPU utilization for five seconds: 54%; one minute: 46%; five minutes: 46%
Core 2: CPU utilization for five seconds: 75%; one minute: 63%; five minutes: 58%
Core 3: CPU utilization for five seconds: 48%; one minute: 49%; five minutes: 57%
PID    T C  TID    Runtime(ms) Invoked uSecs  5Sec      1Min     5Min     TTY   Process
12577  L          2766882     2422952 291    23.52     23.67    23.69    34816 iosd
12577  L 3  12577  1911782     1970561 0      23.34     23.29    23.29    34818 iosd
12577  L 0  14135  694490      3264088 0      0.28      0.34     0.36     0     iosd.fastpath
162    I          2832830     6643    0      93.11     92.55    92.33    0     NGWC L2M
```

Soluzione

Configurare **lo snooping mld ipv6** sugli switch interessati in modo da abilitare globalmente lo **snooping mld ipv6**. Ciò dovrebbe ridurre l'utilizzo della CPU.

```
3850#conf t
Enter configuration commands, one per line. End with CNTL/Z.
3850(config)#ipv6 mld snooping
3850(config)#end
```

Quando lo snooping MLD è abilitato, nel software e nell'hardware viene creata una tabella degli indirizzi multicast IPv6 per VLAN. Lo switch esegue quindi il bridging basato su indirizzo multicast IPv6 nell'hardware, impedendo l'elaborazione di questi pacchetti da parte del software.

Fare clic sul collegamento per ulteriori informazioni sulla configurazione dello snooping MLD

Nelle versioni precedenti di IOS XE, è stato rilevato che la coda della CPU potrebbe bloccarsi a causa di questo problema che impedirebbe a tutti i pacchetti di controllo presenti nella coda di raggiungere la CPU. Questa condizione è stata risolta tramite CSCuo14829 nelle versioni 3.3.3 e 3.6.0 di IOS e successive. Fare riferimento a questo bug per i dettagli.

Switch Catalyst serie 4500

Gli switch Catalyst serie 4500 supportano l'inoltro hardware del traffico multicast IPv6 tramite TCAM (Ternary Content Addressable Memory). Questa condizione viene spiegata in Multicast sugli switch Cisco Catalyst serie 4500E e 4500X

Per quanto riguarda il traffico di rilevamento del listener multicast IPv6, lo switch deve eseguire l'inoltro software (utilizzando risorse CPU). Come spiegato in Configurazione dello snooping IPv6 MLD sugli switch Catalyst 4500, lo snooping MLD può essere abilitato o disabilitato a livello globale o per VLAN. Quando lo snooping MLD è abilitato, nel software viene creata una tabella degli indirizzi MAC multicast IPv6 per VLAN e una tabella degli indirizzi multicast IPv6 per VLAN nel software e nell'hardware. Lo switch esegue quindi il bridging basato su indirizzo multicast IPv6 nell'hardware. Questo è il comportamento previsto sugli switch Catalyst serie 4500.

Per controllare il tipo di pacchetto da reindirizzare alla CPU, è possibile eseguire il comando "**debug platform packet all buffer**" seguito dal comando "**show platform cpu packet buffered**".

```
4500#debug platform packet all buffer
platform packet debugging is on
Cat4500#sh platform cpu packet buffered
Total Received Packets Buffered: 1024
-------------------------------------
Index 0:
33 days 11:42:21:833532 - RxVlan: 214, RxPort: Te1/15
Priority: Normal, Tag: Dot1Q Tag, Event: L2 Router, Flags: 0x40, Size: 90
Eth: Src 44:39:C4:39:5A:4A Dst 33:33:FF:7F:EB:DB Type/Len 0x86DD
Remaining data&colon;
0: 0x60 0x0  0x0  0x0  0x0  0x20 0x0  0x1  0xFE 0x80
10: 0x0  0x0  0x0  0x0  0x0  0x0  0x46 0x39 0xC4 0xFF
20: 0xFE 0x39 0x5A 0x4A 0xFF 0x2  0x0  0x0  0x0  0x0
30: 0x0  0x0  0x0  0x0  0x0  0x1  0xFF 0x7F 0xEB 0xDB
40: 0x3A 0x0  0x5  0x2  0x0  0x0  0x1  0x0  0x83 0x0
```

Questo pacchetto è arrivato sull'interfaccia Tengigabitethernet1/15 sulla vlan 214 dall'indirizzo MAC di origine 44:39:C4:39:5A:4A. Il protocollo 0x86DD è IPv6 e l'indirizzo MAC 33:33:FF:7F:EB:DB è utilizzato in questo caso per i nodi MLD IPv6 multicast.

Soluzione

Sono disponibili due opzioni per correggere l'elevato utilizzo della CPU dovuto a questo traffico.

1. Disabilitare la generazione del traffico di individuazione del listener multicast IPv6 sull'host finale. A tale scopo, è possibile aggiornare i driver NIC o disattivare la funzionalità nel BIOS degli host che inviano pacchetti IPv6. È possibile contattare il fornitore del computer client per disattivare le funzionalità del BIOS o aggiornare i driver NIC.

2. Abilitare Control Plane Policing (CoPP) per eliminare la quantità eccessiva di traffico di individuazione del listener multicast IPv6 puntato alla CPU. Inoltre, poiché questi pacchetti sono costituiti dal limite di hop di un collegamento locale, si prevede che verranno indirizzati alla CPU.

```
ipv6 access-list IPv6-Block
permit ipv6 any any
!
class-map TEST
match access-group name IPv6-Block
!
policy-map ipv6
class TEST
police 32000 conform-action drop exceed-action drop
!
control-plane
service-policy input ipv6
```

Nell'esempio precedente, la quantità di traffico IPv6 gestito dalla CPU viene limitata a 32000 pacchetti al secondo.

Switch Catalyst serie 6500

Gli switch Catalyst 6500 prendono decisioni sull'inoltro nell'hardware che utilizza TCAM e che normalmente non richiede assistenza alla CPU, purché TCAM abbia la voce di inoltro.

Supervisor Engine 720 sugli switch Catalyst 6500 ha due CPU. Una CPU è il Network Management Processor (NMP) o lo Switch Processor (SP). L'altra CPU è la CPU di layer 3, detta Route Processor (RP).

L'utilizzo della CPU da parte dei processi e degli interrupt è elencato nel comando **show process cpu**. Come mostrato di seguito, La CPU causata dagli interrupt è per lo più basata sul traffico. Il traffico con commutazione di interrupt è il traffico che non corrisponde a un processo specifico, ma deve ancora essere inoltrato. L'esempio mostra uno switch Catalyst 6500 con un elevato utilizzo della CPU sull'RP a causa di interrupt.

```
6500#show process cpu
CPU utilization for five seconds: 98%/92%;
one minute: 99%; five minutes: 99% PID Runtime(ms)   Invoked
```

Verificare se un'interfaccia o una VLAN di layer 3 stanno riducendo una quantità elevata di traffico. (la coda di input si interrompe). In tal caso, il traffico potrebbe essere indirizzato al server RP da tale vlan.

```
Vlan19 is up, line protocol is up
 Input queue: 0/75/6303532/0 (size/max/drops/flushes); Total output drops: 0
 Queueing strategy: fifo
 5 minute input rate 19932000 bits/sec, 26424 packets/sec
 5 minute output rate 2662000 bits/sec, 1168 packets/sec
```

Il comando seguente può essere usato per trovare tutti i pacchetti nel buffer della coda di input per l'interfaccia vlan 19.

```
6500#show buffer input-interface vlan 19 packet
```

In alternativa, è possibile usare l'acquisizione NetDR per acquisire il traffico diretto alla CPU su uno switch Catalyst 6500. Questo documento spiega come interpretare i pacchetti acquisiti con l'acquisizione NetDR.

```
------- dump of incoming inband packet -------
interface Vl16, routine mistral_process_rx_packet_inlin, timestamp 03:17:56.380
dbus info: src_vlan 0x10(16), src_indx 0x1001(4097), len 0x5A(90)
  bpdu 0, index_dir 0, flood 1, dont_lrn 0, dest_indx 0x4010(16400)
  E8820000 00100000 10010000 5A080000 0C000418 01000008 00000008 4010417E
mistral hdr: req_token 0x0(0), src_index 0x1001(4097), rx_offset 0x76(118)
  requeue 0, obl_pkt 0, vlan 0x10(16)
destmac 33.33.FF.4A.C3.FD, srcmac C8.CB.B8.29.33.62, protocol 86DD
protocol ipv6: version 6, flow 1610612736, payload 32, nexthdr 0, hoplt 1
class 0, src FE80::CACB:B8FF:FE29:3362, dst FF02::1:FF4A:C3FD
```

Soluzione

Utilizzare una o più delle soluzioni seguenti.

1. Elimina pacchetti multicast IPv6 utilizzando la configurazione seguente.

```
6500(config)#mac-address-table static 3333.FF4A.C3FD vlan <vlan #> drop
```

2. Reindirizzare il traffico multicast IPv6 a un'interfaccia di arresto inutilizzata o di amministrazione (Gi1/22 in questo esempio).

```
6500(config)#mac-address-table 3333.FF4A.C3FD vlan 19 interface Gi1/22
```

3. Utilizzare il VACL (Vlan Access Control List) per eliminare il traffico multicast IPv6.

```
6500(config)#mac access-li extended Multicast_MAC
6500(config-ext-macl)#permit any host 3333.FF4A.C3FD
6500(config-ext-macl)#exit
6500(config)#vlan access-map block-ipv6 10
6500(config-access-map)#action drop
6500(config-access-map)#match mac address Multicast_MAC
6500(config-access-map)#exit
6500(config-access-map)#vlan access-map block-ipv6 20
6500(config-access-map)#action forward
6500(config-access-map)#exit
6500(config)#vlan filter block-ipv6 vlan-list <vlan #>
```

4. Disabilitare lo snooping MLD IPv6.

```
6500(config)#no ipv6 mld snoopin
```

5. Elimina traffico multicast IPv6 tramite Control Plane Policing (CoPP)

```
6500(config)#ipv6 access-list test
6500(config-ipv6-acl)#permit ipv6 any any
6500(config-ipv6-acl)#exit

6500(config)#class-map TEST
6500(config-cmap)#match access-group name test
6500(config-cmap)#exit

6500(config)#policy-map ipv6
6500(config-pmap)#class TEST
6500(config-pmap-c)#police 320000 conform-action drop exceed-action drop
6500(config-pmap-c)#exit

6500(config)#control-plane
6500(config-cp)#service-policy in ipv6
6500(config-cp)#exit
```

6. Controllo della tempesta sulle interfacce in entrata. storm-control monitora i livelli di traffico in entrata su un intervallo di 1 secondo e durante questo intervallo confronta il livello di traffico con il livello di controllo configurato per l'uragano. Il livello di controllo della tempesta sul traffico è una percentuale della larghezza di banda totale disponibile sulla porta. Ogni porta dispone di un singolo livello di controllo Traffic Storm utilizzato per tutti i tipi di traffico (broadcast, multicast e unicast).

```
6500(config)#interface Gi2/22
6500(config-if)#storm-control multicast level 10
```

7. Se la CPU è alta sull'SP (Switch Processor), applicare la soluzione indicata di seguito.

```
6500(config)#mls rate-limit ipv6 mld 10 1
```

Se non è possibile determinare il motivo della mancata corrispondenza delle informazioni fornite in questo documento, aprire una richiesta al servizio TAC per ulteriori informazioni.