Uso elevado de la CPU en switches Catalyst debido al tráfico de multidifusión IPv6

Contenido

Introducción

Este documento describe el uso elevado de la CPU en varias plataformas Catalyst debido a la inundación de paquetes IPV6 Multicast Listener Discovery y maneras de mitigar este problema.

Prerequisites

No hay requisitos previos.

Requirements

No hay requisitos específicos para este documento.

Componentes Utilizados

La información de este documento se basa en los switches Catalyst de Cisco serie 6500, los switches Catalyst serie 4500 y los switches Catalyst serie 3850.

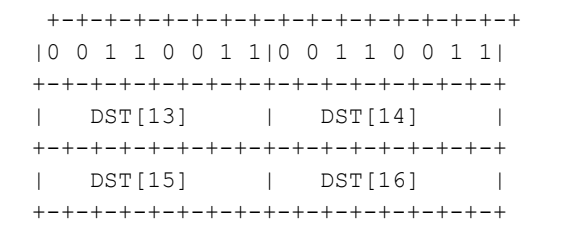
The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration.

Problema

Se puede ver una alta utilización de la CPU en algunas plataformas Cisco Catalyst debido a que

el tráfico de multidifusión IPv6 con dirección MAC en el rango 333.xxxx.xxxx se impulsa a la CPU.

Según RFC7042, todos los identificadores de multidifusión MAC-48 tienen como prefijo "33-33" (es decir, los 2**32 identificadores de MAC de multidifusión en el rango de 33-33-00-00-00-00 a 33-33-FF-FF-FF-FF) se utilizan como se especifica en [2. 464] para multidifusión IPv6. Un paquete IPv6 con una dirección de destino de multidifusión DST, que consta de los dieciséis octetos DST[1] a DST[16], se transmite a la dirección de multidifusión Ethernet cuyos dos primeros octetos son el valor hexadecimal 333 y cuyos cuatro últimos octetos son los cuatro últimos octetos de DST, como se muestra en la figura 1.



(Figure 1: Address Mapping - Multicast)

En algunas ocasiones se ha observado que cuando los dispositivos host que utilizan una tarjeta NIC determinada pasan al modo inactivo, inundan el tráfico de multidifusión IPv6. Este problema no se limita a un proveedor de host concreto, aunque se ha visto que algunos chipsets muestran este comportamiento más a menudo que otros.

Solución de problemas y solución

Puede utilizar los siguientes procedimientos para averiguar si el switch Catalyst que observa un uso elevado de la CPU se ve afectado por este problema e implementar las soluciones correspondientes.

Catalyst 3850 Series Switches

En los switches Catalyst 3850, el proceso NGWC L2M utiliza la CPU para procesar los paquetes IPv6. Cuando se inhabilita el snooping de descubrimiento de receptor multidifusión (MLD) en el switch, el paquete de unión/abandono MLD se inunda en todos los puertos miembro. Y, si hay muchos paquetes entrantes MLD de unión/abandono, este proceso consumirá más ciclos de CPU para enviar los paquetes en todos los puertos miembro. Se ha visto que cuando ciertas máquinas host pasan al modo suspendido, pueden enviar varios miles de paquetes/seg de tráfico de IGMPv6 MLD.

```
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
```

Solución

Configure la **indagación de mld ipv6** en los switches afectados para habilitar globalmente la **indagación de mld ipv6**. Esto debería reducir el uso de la CPU.

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

Cuando se habilita la indagación MLD, se construye una tabla de direcciones multidifusión IPv6 por VLAN en software y hardware. A continuación, el switch realiza el bridging basado en dirección multicast IPv6 en el hardware, lo que evita que estos paquetes sean procesados por el software.

Haga clic en el enlace para obtener más información sobre la configuración de MLD Snooping

En las versiones anteriores de IOS XE, se encontró que la cola de la CPU podía atascarse debido a este problema que evitaría que todos los paquetes de control en esa cola fueran a la CPU. Esto se corrigió a través de CSCuo14829 en las versiones 3.3.3 y 3.6.0 y posteriores del IOS. Consulte este error para obtener más información.

Catalyst 4500 Series Switches

Los switches Catalyst de la serie 4500 admiten el reenvío de hardware del tráfico de multidifusión IPv6 utilizando la memoria direccionable de contenido ternario (TCAM). Esto se explica en Multicast en Cisco Catalyst 4500E y 4500X Series Switches

Cuando se trata del tráfico de detección de receptor multidifusión IPv6, el switch debe realizar el reenvío de software (mediante recursos de CPU). Como se explica en Configuración de la Indagación MLD de IPv6 en los switches Catalyst 4500, la indagación MLD se puede habilitar o inhabilitar globalmente o por VLAN. Cuando se habilita la indagación MLD, se construye una tabla de direcciones MAC multicast por VLAN IPv6 en software y una tabla de direcciones multicast por VLAN IPv6 se construye en software y hardware. A continuación, el switch realiza el bridging basado en dirección multicast IPv6 en el hardware. Este es el comportamiento esperado en los switches Catalyst de la serie 4500.

Para verificar el tipo de paquete que se impulsa a la CPU, podemos ejecutar "**debug platform packet all buffer**" seguido del 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
```

Este paquete llegó a la interfaz Tengigabitethernet1/15 en vlan 214 desde la dirección mac de origen 44:39:C4:39:5A:4A. El protocolo 0x86DD es IPv6 y Dst MAC 33:33:FF:7F:EB:DB se está utilizando para los nodos MLD IPv6 de multidifusión en este caso.

Solución

Tenemos dos opciones para solucionar el uso elevado de la CPU debido a este tráfico.

1. Inhabilite la generación del tráfico de detección de receptor multidifusión IPv6 en el host final. Esto puede hacerse actualizando los controladores NIC o desactivando la función en el BIOS de los hosts que envían paquetes IPv6. Puede ponerse en contacto con el proveedor de la máquina cliente que puede ayudarle a desactivar la función en el BIOS o actualizar los controladores NIC.

2. Habilite Control Plane Policing (CoPP) para descartar la cantidad excesiva de tráfico IPv6 Multicast Listener Discovery que se impulsa a la CPU. Y, estos paquetes son el límite de saltos de un link local, por lo que se espera que estos paquetes sean impulsados a la 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
```

En el ejemplo anterior, estamos limitando la cantidad de tráfico IPv6 que maneja la CPU a 32000 paquetes por segundo.

Catalyst 6500 Series Switches

Los switches Catalyst 6500 toman decisiones de reenvío en hardware usando TCAM que normalmente no necesita asistencia de la CPU mientras TCAM tenga la entrada de reenvío.

Supervisor Enginet 720 en los switches Catalyst 6500 tiene dos CPU. Una CPU es el Procesador de administración de red (NMP) o el Procesador de switch (SP). La otra CPU es la CPU de Capa 3, que se denomina Procesador de ruta (RP).

El proceso y la utilización de la CPU de interrupción se enumeran en el comando **show process cpu**. Como se muestra a continuación, High La CPU causada por interrupciones se basa principalmente en el tráfico. El tráfico conmutado de interrupción, es el tráfico que no coincide con un proceso específico, pero que aún debe reenviarse. El siguiente ejemplo muestra un switch Catalyst 6500 con uso elevado de CPU en RP debido a interrupciones.

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

Verifique si alguna interfaz o VLAN de Capa 3 está descartando una gran cantidad de tráfico. (Cola de entrada descartada). Si es así, el tráfico puede estar siendo impulsado al RP desde esa 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
```

El siguiente comando se puede utilizar para encontrar todos los paquetes en el búfer de cola de entrada para la interfaz vlan 19.

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

Alternativamente, puede utilizar la captura de NetDR para capturar el tráfico que va a la CPU en un switch Catalyst 6500. Este documento explica cómo interpretar los paquetes capturados mediante la captura de 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
```

Solución

Utilice una o varias de las siguientes soluciones.

1. Elimine los paquetes de multidifusión IPv6 mediante la siguiente configuración.

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

2. Redireccione el tráfico de multidifusión IPv6 a una interfaz de apagado no utilizada o de administración (Gi1/22 en este ejemplo).

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

3. Utilice Vlan Access Control List (VACL) para descartar el tráfico de multidifusión 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. Inhabilite la indagación MLD IPv6.

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

5. Eliminación del tráfico de multidifusión IPv6 mediante 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. Utilice el control de tormentas en las interfaces de ingreso. El control de tormentas monitorea los niveles de tráfico entrante durante un intervalo de 1 segundo y durante este intervalo compara el nivel de tráfico con el nivel de control de tormentas de tráfico configurado. El nivel de control de tormentas de tráfico es un porcentaje del ancho de banda total disponible del puerto. Cada puerto tiene un único nivel de control de tormentas de tráfico que se utiliza para todos los tipos de tráfico (difusión, multidifusión y unidifusión).

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

7. En caso de que la CPU sea Alta en SP (Procesador del switch), aplique la siguiente solución alternativa.

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

Si no puede determinar el motivo basándose en la información proporcionada en este documento, abra una solicitud de servicio del TAC para investigar más a fondo.