Alta utilização da CPU em Switches Catalyst devido ao tráfego multicast IPv6

Contents

[Introduction](#)
[Prerequisites](#)
[Requirements](#)
[Componentes Utilizados](#)
[Problema](#)
[Solução de problemas e solução](#)
[Catalyst 3850 Series Switches](#)
[Solução](#)
[Catalyst 4500 Series Switches](#)
[Solução](#)
[Catalyst 6500 Series Switches](#)
[Solução](#)
[Discussões relacionadas da comunidade de suporte da Cisco](#)

Introduction

Este documento descreve a alta utilização da CPU em várias plataformas Catalyst devido à inundação de pacotes IPV6 Multicast Listener Discovery e maneiras de atenuar esse problema.

Prerequisites

Não há pré-requisitos.

Requirements

Não existem requisitos específicos para este documento.

Componentes Utilizados

As informações neste documento são baseadas nos Cisco Catalyst 6500 Series Switches, Catalyst 4500 Series Switches e Catalyst 3850 Series Switches.

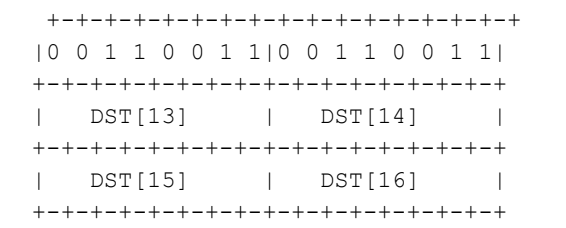
The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration.

Problema

A alta utilização da CPU pode ser vista em algumas plataformas Cisco Catalyst devido ao tráfego

multicast IPv6 com endereço MAC no intervalo 333.xxxx.xxxx sendo direcionado para a CPU.

Conforme o RFC7042, todos os identificadores multicast MAC-48 prefixos "33-33" (ou seja, os identificadores MAC multicast 2**32 no intervalo de 33-33-00-00-00-00 a 33-33-FF-FF-FF-FF) são usados como especificado em [RFC2464] para multicast IPv6. Um pacote IPv6 com um endereço de destino de multicast DST, consistindo dos dezesseis octetos DST[1] a DST[16], é transmitido ao endereço de multicast Ethernet cujos dois primeiros octetos são o valor 333 hexadecimal e cujos últimos quatro octetos são os últimos quatro octetos do DST, como mostrado na Figura 1.



(Figure 1: Address Mapping - Multicast)

Em algumas ocasiões, observou-se que, quando os dispositivos hosts que usam uma determinada placa de rede entram no modo de espera, inundam o tráfego multicast IPv6. Esse problema não se limita a um fornecedor específico de host, embora determinados chipsets tenham exibido esse comportamento com mais frequência do que outros.

Solução de problemas e solução

Você pode usar os procedimentos a seguir para descobrir se o seu switch Catalyst que observa alta utilização da CPU é afetado por esse problema e implementar as respectivas soluções.

Catalyst 3850 Series Switches

Nos switches Catalyst 3850, o processo NGWC L2M usa a CPU para processar pacotes IPv6. Quando a espionagem de Multicast Listener Discovery (MLD) é desabilitada no switch, a junção/saída de MLD é inundada em todas as portas membro. E, se houver muitos pacotes de entrada/saída de MLD, esse processo consumirá mais ciclos de CPU para enviar os pacotes em todas as portas membro. Foi observado que, quando certas máquinas host entram no modo de espera, elas podem enviar vários milhares de pacotes/s de tráfego 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
```

Solução

Configure **ipv6 mld snooping** nos switches afetados para habilitar globalmente **a espionagem de mld ipv6**. Isso deve reduzir a utilização da CPU.

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

Quando o rastreamento de MLD é ativado, uma tabela de endereços multicast IPv6 por VLAN é construída em software e hardware. Em seguida, o switch executa o bridging baseado em endereço multicast IPv6 no hardware, o que impede que esses pacotes sejam processados pelo software.

Clique no link para obter mais informações sobre como configurar rastreamento MLD

Em versões anteriores do IOS XE, foi descoberto que a fila da CPU poderia ficar presa devido a esse problema que impediria todos os pacotes de controle nessa fila de irem para a CPU. Isso foi corrigido através do CSCuo14829 nas versões 3.3.3 e 3.6.0 do IOS e posteriores. Consulte este bug para obter detalhes.

Catalyst 4500 Series Switches

Os switches da série Catalyst 4500 suportam encaminhamento de hardware de tráfego multicast IPv6 usando TCAM (Ternary Content Addressable Memory). Isso é explicado em multicast nos switches Cisco Catalyst 4500E e 4500X Series

Quando se trata de tráfego de descoberta de ouvinte de multicast IPv6, o switch precisa executar o encaminhamento de software (usando recursos da CPU). Conforme explicado em Configurando o rastreamento MLD IPv6 em Switches Catalyst 4500, o rastreamento MLD pode ser ativado ou desativado globalmente ou por VLAN. Quando a espionagem de MLD é habilitada, uma tabela de endereços MAC multicast IPv6 por VLAN é construída no software e uma tabela de endereços multicast IPv6 por VLAN é construída no software e no hardware. Em seguida, o switch executa o bridging baseado em endereço multicast IPv6 no hardware. Esse é o comportamento esperado nos switches da série Catalyst 4500.

Para verificar o tipo de pacote que está sendo direcionado para a CPU, podemos executar **"debug platform packet all buffer"** seguido do 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
```

Esse pacote chegou à interface Tengigabitethernet1/15 na vlan 214 do endereço mac de origem

44:39:C4:39:5A:4A. O protocolo 0x86DD é IPv6 e o Dst MAC 33:33:FF:7F:EB:DB está sendo usado para nós MLD IPv6 Multicast neste caso.

Solução

Temos duas opções para corrigir a alta utilização da CPU devido a esse tráfego.

1. Desative a geração de tráfego de descoberta de ouvinte multicast IPv6 no host final. Isso pode ser feito atualizando os drivers da placa de rede ou desabilitando o recurso no BIOS dos hosts que enviam pacotes IPv6. Você pode entrar em contato com o fornecedor da máquina cliente, que pode ajudar a desabilitar o recurso no BIOS ou atualizar os drivers da placa de rede.

2. Ative o Control Plane Policing (CoPP) para descartar a quantidade excessiva de tráfego IPv6 Multicast Listener Discovery que está sendo direcionado para a CPU. E esses pacotes são o limite de saltos de um link local, portanto, é esperado que o comportamento desses pacotes seja direcionado para a 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
```

No exemplo acima, estamos limitando a quantidade de tráfego IPv6 que é tratado pela CPU para 32.000 pacotes por segundo.

Catalyst 6500 Series Switches

Os Switches Catalyst 6500 tomam decisões de encaminhamento em hardware usando TCAM, que normalmente não precisa de assistência da CPU, desde que a TCAM tenha a entrada de encaminhamento.

O Supervisor Enginet 720 em Catalyst 6500 Switches tem duas CPUs. Uma CPU é o Network Management Processor (NMP) ou o Switch Processor (SP). A outra CPU é a Camada 3, chamada de Route Processor (RP).

A utilização da CPU do processo e da interrupção está listada no comando **show process cpu**. Como mostrado abaixo, alto A CPU causada por interrupções é principalmente baseada em tráfego. O tráfego comutado por interrupção é o tráfego que não corresponde a um processo específico, mas que ainda precisa ser encaminhado. O exemplo a seguir mostra um switch Catalyst 6500 com alta utilização da CPU no RP devido a interrupções.

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

Verifique se alguma interface ou Vlan de Camada 3 está descartando grande quantidade de tráfego. (A fila de entrada cai). Se sim, o tráfego pode estar sendo direcionado para o RP a partir dessa 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
```

O seguinte comando pode ser usado para localizar todos os pacotes no buffer de fila de entrada para a interface vlan 19.

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

Como alternativa, você pode usar a captura NetDR para capturar o tráfego que vai para a CPU em um switch Catalyst 6500. Este documento explica como interpretar os pacotes capturados usando a captura 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
```

Solução

Use uma ou mais das soluções abaixo.

1. Descarte os pacotes multicast IPv6 usando a seguinte configuração.

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

2. Redirecione o tráfego multicast IPv6 para uma interface de desligamento não utilizada ou admin (Gi1/22 neste exemplo).

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

3. Use a VACL (Lista de Controle de Acesso) para descartar o tráfego 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. Desative a espionagem MLD IPv6.

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

5. Descartar o tráfego multicast IPv6 usando a Política de plano de controle (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. Use o controle de tempestade em interfaces de entrada. o controle de tempestade monitora os níveis de tráfego de entrada em um intervalo de 1 segundo e, durante esse intervalo, compara o nível de tráfego com o nível de controle de tempestade de tráfego configurado. O nível de controle de tempestade de tráfego é uma porcentagem da largura de banda total disponível da porta. Cada porta tem um único nível de controle de tempestade de tráfego que é usado para todos os tipos de tráfego (broadcast, multicast e unicast).

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

7. Caso a CPU seja alta no SP (Switch Processor), aplique a solução alternativa abaixo.

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

Se você não puder determinar o motivo com base nas informações fornecidas neste documento, abra uma solicitação de serviço do TAC para investigar mais.