Implementar as práticas recomendadas de SSDP nos switches Catalyst 9000 Series

Contents

Introduction

Este documento descreve as configurações de práticas recomendadas projetadas para descartar ou limitar os pacotes do protocolo SSDP nos switches da série Catalyst 9000.

Prerequisites

Requirements

A Cisco recomenda que você tenha conhecimento destes tópicos:

- Operação de Protocol Independent Multicast (PIM)
- Como o SSDP é usado específico para seu ambiente

Componentes Utilizados

As informações neste documento são baseadas nestas versões de software e hardware:

- Cisco Catalyst 9200
- Cisco Catalyst 9300
- Cisco Catalyst 9400
- Cisco Catalyst 9500
- Cisco Catalyst 9600

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. Se a rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

Compreender os riscos do SSDP em ambientes corporativos

Em geral, dispositivos de usuário final, como laptops e telefones móveis, anunciam

automaticamente seus recursos UPnP (Universal Plug-and-Play) que usam o protocolo SSDP. Os clientes enviam um pacote de anúncio multicast ao endereço IP 239.255.255.250. Esses anúncios são frequentemente enviados com um Time to Live (TTL) de 1 e não vão além da sub-rede local dos hosts que geraram o pacote multicast. Para receber os anúncios de outros dispositivos na rede, os endpoints também enviam um Relatório de Associação IGMP para o endereço 239.255.255.250, que informa à rede que o tráfego multicast enviado a esse endereço IP de qualquer outra origem multicast também deve ser encaminhado a esse cliente.

Em ambientes corporativos que contêm centenas ou milhares de endpoints, todos atuando como origem e como receptor interessado desse grupo, essa atividade do cliente pode facilmente sobrecarregar os dispositivos de rede se não for verificada e pode causar interrupções quando os recursos da rede estiverem esgotados.

Essa exaustão acontece principalmente de duas maneiras:

1. Esgotamento de recursos de hardware que aciona falhas de protocolo secundário
2. Esgotamento da largura de banda da interface e da plataforma do SSDP usado como um ataque de negação de serviço distribuído (DDoS).

Embora não seja discutido em detalhes neste documento, deve-se observar que, devido à natureza aberta do SSDP, é possível para um invasor enviar um pacote criado para um grupo de clientes com este serviço habilitado, a fim de disparar uma resposta grande a ser enviada para um ou um grupo de hosts de destino. A grande quantidade de estado de interface de saída que é criada também significa que a capacidade de desempenho do switch pode ser significativamente sobrecarregada a partir de uma pequena quantidade de tráfego multicast, já que o switch é obrigado a fazer uma cópia de cada quadro para cada interface de saída no Circuito Integrado Específico da Aplicação (ASIC). A interface de saída lista que as interfaces de número 20 ou mais correm um risco maior de problemas de capacidade e perda de pacotes.

Sintomas de esgotamento de recursos de hardware

Os switches da série Catalyst 9000 imprimem syslogs que mencionam "fman_fp_image" ou "FMFP" quando os recursos estão esgotados. Alguns ou todos esses erros podem ser impressos quando o switch está esgotado em recursos e precisa ser investigado mais a fundo.

Estes são alguns dos erros mais comuns vistos durante o esgotamento de recursos, mas não é uma lista abrangente.

Figura 1: Exemplo dos erros mais comuns impressos que são evidência de esgotamento de recursos em um switch

```
%FMFP-3-OBJ_DWNLD_TO_DP_STUCK: R0/0: fman_fp_image: AOM download to Data Plane is stuck for more
than 1800 seconds for <object details>
%FMFP-3-OBJ_DWNLD_TO_DP_RESUME: R0/0: fman_fp_image: AOM download of objects to Data Plane is
back to normal
%FMFP_QOS-6-QOS_STATS_STALLED: R0/0: fman_fp_image: statistics stalled
%FMFP-3-OBJ_DWNLD_TO_DP_FAILED: R0/0: fman_fp_image: adj <hex>, Flags None download to DP failed
%FMFP-3-OBJ_DWNLD_TO_DP_FAILED: R0/0: fman_fp_image: adj <hex>, Flags Midchain download to DP
failed
%FED_L3M_ERRMSG-3-RSRC_ERR: Switch <num> R0/0: fed: Failed to allocate hardware resource for
group <address> - rc:<number or error>
%FED_L3_ERRMSG-3-RSRC_ERR: Chassis <num> R0/0: fed: Failed to allocate hardware resource for adj
```

```
entry due to hardware resource exhaustion - rc:<number or error>
```

Verifique o esgotamento de recursos de hardware causado pelo SSDP

Todos os switches da série Catalyst 9000 utilizam ASICs especiais para executar a maioria do roteamento de pacotes em alto throughput. Esses ASICs aproveitam diferentes tabelas e recursos internos que são finitos em sua capacidade. Como os clientes SSDP agem como origens e receptores para um grupo multicast comum, o hardware deve usar esses recursos limitados para programar um caminho no hardware para os pacotes seguirem, mesmo que esses pacotes nunca cheguem ou sejam descartados por outros motivos (TTL 1). Quando os recursos de hardware estiverem esgotados, nenhuma atualização ou adição nova para qualquer grupo, independentemente de sua relação com o SSDP, poderá ser instalada. Um grande número de atualizações de SSDP não instaladas (rotatividade de estado) também pode ser enfileirado no software, isso também pode fazer com que as atualizações de hardware para tráfego não multicast sejam interrompidas ou falhem, o que afeta o tráfego do usuário e causa interrupções da rede.

Este documento só é relevante se a sua rede estiver configurada com PIM e tiver estado multicast da camada 3 para o endereço de grupo conhecido SSDP. Para verificar esses critérios, execute o comando "**show ip mroute 239.255.255.250**" (adicione instruções vrf, se necessário). O grupo 239.255.255.250 é específico do protocolo SSDP.

Se a saída do comando contiver um grande número de interfaces de saída e/ou tiver um grande número de origens exclusivas para esse grupo específico, isso indica que o sistema e a rede estão vulneráveis a interrupções causadas pelo SSDP. Quanto maior o número de interfaces de saída e origens exclusivas, maiores as chances de que isso possa causar impacto no serviço.

Figura 2: Exemplo de saída de "show ip mroute 239.255.255.250" com o SSDP ativo na rede.

```
Switch#show ip mroute 239.255.255.250
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       G - Received BGP C-Mroute, g - Sent BGP C-Mroute,
       N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed,
       Q - Received BGP S-A Route, q - Sent BGP S-A Route,
       V - RD & Vector, v - Vector, p - PIM Joins on route,
       x - VxLAN group
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.255.255.250), 00:08:35/stopped, RP 10.0.0.1, flags: SJC
  Incoming interface: GigabitEthernet0/0/1.40, RPF nbr 10.0.0.1
  Outgoing interface list:
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:08:35/00:02:40
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:08:35/00:02:38
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:08:35/00:02:39
```

```
(10.1.1.2, 239.255.255.250), 00:01:40/00:01:19, flags: T
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    GigabitEthernet0/0/1.40, Forward/Sparse, 00:01:40/00:01:40, A
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:01:40/00:02:39
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:01:40/00:02:38
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:01:40/00:02:40

(10.1.1.3, 239.255.255.250), 00:02:03/00:00:56, flags: JT
  Incoming interface: GigabitEthernet0/0/1.40, RPF nbr 10.1.1.1
  Outgoing interface list:
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:02:03/00:02:39
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:02:03/00:02:38
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:02:03/00:02:40

(10.1.1.4, 239.255.255.250), 00:08:35/00:02:32, flags: T
  Incoming interface: GigabitEthernet0/0/1.40, RPF nbr 10.1.1.1
  Outgoing interface list:
    GigabitEthernet0/0/1.100, Forward/Sparse, 00:08:35/00:02:39
    GigabitEthernet0/0/1.102, Forward/Sparse, 00:08:35/00:02:38
    GigabitEthernet0/0/1.101, Forward/Sparse, 00:08:35/00:02:40, A
```

A menos que o SSDP seja usado para uma finalidade específica, espera-se que essa saída esteja vazia ou que tenha um número baixo de interfaces de saída e/ou que tenha um número baixo de fontes exclusivas para evitar o esgotamento de recursos e possíveis impactos no serviço.

Se um grande número de grupos multicast for visto, o comando "**show platform software object-manager fp ative statistics**" ou "**show platform software object-manager fp switch ative statistics**" poderá ser usado para informar se um recurso de hardware foi esgotado.

> **Note**: Esse comando não é específico para a exaustão de recursos disparada pelo tráfego multicast; outros problemas podem fazer com que esses valores sejam diferentes de zero.

Figura 3: Saída de **"show platform software object-manager fp active statistics"**em estado de problema

```
Switch#show platform software object-manager fp active statistics
Forwarding Manager Asynchronous Object Manager Statistics
Object update: Pending-issue: 109058, Pending-acknowledgement: 76928   <-- Pending-issue is very
high, this
Batch begin:   Pending-issue: 0, Pending-acknowledgement: 0                  is not expected.
Batch end:     Pending-issue: 0, Pending-acknowledgement: 0
Command:       Pending-acknowledgement: 0
Total-objects: 304085
Stale-objects: 0
Resolve-objects: 0
Childless-delete-objects: 530
Error-objects: 1098

Paused-types: 127
```

A saída da figura 3 demonstra os sintomas de um switch com esgotamento de recursos. Há várias linhas de saída de comando que não são esperadas durante a operação normal:

- Emissão pendente: Espera-se que seja zero ou próximo a ele. Se esse valor permanecer grande, diferente de zero em várias iterações do comando, isso é um sinal de esgotamento

de recursos
- Confirmação pendente: Espera-se que seja zero ou próximo a ele. Se esse valor permanecer grande, diferente de zero em várias iterações do comando, isso é um sinal de esgotamento de recursos
- Objetos de exclusão sem filhos: Espera-se que seja zero ou próximo a ele. Valores de 10+ não são esperados.
- Objetos de erro: Espera-se que seja zero ou próximo a ele. Valores de 10+ não são esperados.

Consistentemente, em um estado onde há um grande número de contadores de "pendências" ou "confirmações pendentes" aumenta o risco de o hardware se tornar mal programado. O hardware programado incorretamente é uma fonte comum de interrupções para tráfego unicast e multicast.

O comando "**show platform hardware fed switch active fwd-asic resource utilization**" or in some models "**show platform hardware fed active fwd-asic resource utilization**" pode ser usado para examinar alguns dos recursos finitos em uso nos ASICs e determinar se um recurso interno foi esgotado:

Figura 4: Exemplo de saída de"show platform hardware fed active fwd-asic resource utilization"com um recurso quase esgotado.

```
Switch#show platform hardware fed active fwd-asic resource utilization
Resource Info for ASIC Instance: 0
Resource Name           Allocated Free
----------------------------------------
RSC_DI                     3822       38076
RSC_FAST_DI                   0         192
RSC_RIET_0                    1        1024
RSC_RIET_1                    0         512
RSC_RIET_2                    0         512
RSC_RIET_3                    0         512
RSC_RIET_4                    0         512
RSC_RIET_5                    0         512
RSC_RIET_6                    0         256
RSC_RIET_7                    0         255
RSC_VLAN_LE                 116        3976
RSC_L3IF_LE                 116        3907
RIM_RSC_DGT                   1         255
RSC_VPN_PREFIX_ID             1       32768
RSC_LABEL_STACK_ID            1       65536
RSC_RI                     7358       82730
RSC_LI_RI                     0         129
RSC_PORT_LE_RI                0        2048
RSC_PORT_LE                   0        1827
RSC_RI_REP                10635      120437
RSC_SI                    11842      119072
RSC_SI_IND                    1         255
RSC_SI_STATS               3550       45602
RSC_RCP1_FID                  1        1023
RSC_RCP2_FID                  1        1023
RSC_RCP3_FID                  1        1023
RSC_RCP4_FID                  1        1023
RSC_LV1_ECR                   1          63
RSC_LV2_ECR                   3         253
RSC_ENH_ECR                   1           0
RSC_RPF_MATCH                12        1012
RSC_PLC                       1        2047
RSC_PLC_PF                    1         255
```

```
RSC_MTU_INDEX                6         250
RSC_EGR_REDIRECT_INDEX       2        2046
RSC_RIL_INDEX 131065 7 <-- Free entries extremely low, this is not expected.
RSC_SIF                      1        1023
RSC_GROUP_LE                 1        1023
RSC_RI_REP_LOCAL             1           0
RSC_EXT_SI                 512       65024
```

Na figura 4, o valor de "RSC_RIL_INDEX" mostra que há 131065 entradas em uso e apenas 7 estão livres. Esse recurso é consumido por um grande número de grupos SSDP exclusivos. Embora não sejam específicos do SSDP, os recursos que têm um número baixo de entradas livres e um número alto de entradas alocadas são sinais de que o switch está próximo de um problema de capacidade e devem ser investigados.

O comando "**show platform hardware fed switch active fwd-asic resource tcam utilization**" or on some models "**show platform hardware fed active fwd-asic resource tcam utilization**" pode ser usado para examinar uma divisão da utilização por ASIC por recurso. Outra assinatura possível do esgotamento do SSDP é a coluna "Used Values" para "L3 Multicast entries" para se aproximar ou ficar no "Max Values".

Figura 5: Exemplo de saída de"show platform hardware fed active fwd-asic resource tcam utilization"em operação normal

```
Switch#show platform hardware fed active fwd-asic resource tcam utilization
CAM Utilization for ASIC  [0]
 Table                                              Max Values        Used Values
 ------------------------------------------------------------------------------
 Unicast MAC addresses                               32768/768       6160/21
 L3 Multicast entries                                32768/768       3544/8      <-- Normal
Utilization, not near Max Values
 L2 Multicast entries                                 2304            181        <-- Normal
Utilization, not near Max Values
 Directly or indirectly connected routes            212992/1536     11903/39
 Input Ipv4 QoS Access Control Entries                5632            17
 Input Non Ipv4 QoS Access Control Entries            2560            36
 Output Ipv4 QoS Access Control Entries               6144            13
 Output Non Ipv4 QoS Access Control Entries           2048            27
 Input Ipv4 Security Access Control Entries           7168            12
 Input Non Ipv4 Security Access Control Entries       5120            76
 Output Ipv4 Security Access Control Entries          7168            11
 Output Non Ipv4 Security Access Control Entries      8192            27
 Ingress Netflow ACEs                                 1024             8
 Policy Based Routing ACEs                            3072            20
 Egress Netflow ACEs                                  1024             8
 Flow SPAN ACEs                                        512             5
 Flow Egress SPAN ACEs                                 512             8
 Control Plane Entries                                1024           235
 Tunnels                                              2816            26
 Lisp Instance Mapping Entries                         512             3
 Input Security Associations                           512             4
 SGT_DGT                                             32768/768         0/1
 CLIENT_LE                                            8192/512         0/0
 INPUT_GROUP_LE                                       1024             0
 OUTPUT_GROUP_LE                                      1024             0
```

Macsec SPD	256	2

Evite o esgotamento de recursos causado pelo SSDP

Para interromper o esgotamento de recursos, o tráfego SSDP deve ser interrompido antes da primeira criação de estado de salto L3 e multicast. A solução mais rápida é usar uma ACL (Access Control List, lista de controle de acesso) IPv4 aplicada à entrada de todas as interfaces L3 configuradas com PIM que veem esse tráfego. Verifique com o comando "**show ip mroute 239.255.255.250**" e examine a "Interface de Entrada" de cada grupo. Isso indica de qual interface L3 a origem do tráfego é e lembre-se de que pode haver mais de uma interface de origem exclusiva. Este exemplo de configuração permite que o SSDP funcione na camada 2 e permite que hosts L2 adjacentes descubram serviços PNP, mas impede que anúncios de clientes sejam encaminhados através dos limites L3 e impede a criação de estado multicast L3 em qualquer roteador ou switch multicast.

Configure uma ACL estendida:

ip access-list extended BLOCK_SSDP remark Block SSDP **deny ip any host 239.255.255.250 <-- Deny SSDP** permit ip any any **<-- Permit any other group**

Configure em cada interface L3 e aplique a ACL na direção de entrada:

```
Switch#configure terminal
Switch(config)#interface vlan100
Switch(config-if)#ip access-group BLOCK_SSDP in
Switch(config-if)#end
```