在Catalyst 9000系列交換機上實施SSDP最佳實踐

目錄

簡介

本文說明在Catalyst 9000系列交換器上捨棄或限制簡易服務探索通訊協定(SSDP)封包的最佳實踐。

必要條件

需求

思科建議您瞭解以下主題：

- 通訊協定無關多點傳送(PIM)作業
- 如何針對您的環境使用SSDP

採用元件

本文中的資訊係根據以下軟體和硬體版本：

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

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

瞭解企業環境中的SSDP風險

一般情況下，諸如筆記型電腦和行動電話等終端使用者裝置會自動通告其使用SSDP協定的通用即插即用(UPnP)功能。客戶端向IP地址239.255.255.250傳送組播通告資料包。這些通告通常以1的生存時間(TTL)傳送，並且不會超出生成組播資料包的主機的本地子網。要接收網路上其他裝置的通告，終端還會將IGMP成員身份報告傳送到239.255.255.250地址，該地址告訴網路，從任何其他組播源傳送到此IP地址的組播流量也必須轉發到此客戶端。

在包含成百上千個終端的企業環境中，這些終端既是源裝置，又是此組的意向接收者，如果不加以檢查，此客戶端活動可能輕易地淹沒網路裝置，並且一旦網路資源耗盡，就會導致網路故障。

這種耗盡主要有兩種方式：

1. 觸發輔助協定故障的硬體資源耗盡
2. SSDP的介面和平台頻寬耗盡被用作分散式拒絕服務(DDoS)攻擊。

雖然本檔案未詳細討論，但必須注意的是，由於SSDP的開放性質，攻擊者可能將精心編制的資料包傳送到啟用此服務的客戶端組，以觸發將大型響應傳送到一個或一組目標主機。建立的大量傳出介面狀態還意味著交換機效能容量可以從少量組播流量中受到顯著影響，因為交換機需要為應用專用積體電路(ASIC)中的每個傳出介面製作每個幀的一個副本。傳出介面列出了數量為20個或更多介面運行容量問題和丟包風險較高的介面。

硬體資源耗盡的症狀

Catalyst 9000系列交換器會列印在資源耗盡時提到「fman_fp_image」或「FMFP」的系統日誌。當交換機已經消耗完資源並需要進一步調查時，可以列印出其中的一些或所有錯誤。

這些是資源耗盡期間出現的一些較常見錯誤，但並非完整的清單。

圖1：打印的表明交換機上資源耗盡的最常見錯誤示例

```
%FMFP-3-OBJ_DWNLD_TO_DP_STUCK: R0/0: fman_fp_image: AOM download to Data Plane is stuck for more than 1
%FMFP-3-OBJ_DWNLD_TO_DP_RESUME: R0/0: fman_fp_image: AOM download of objects to Data Plane is back to n
%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 <add
%FED_L3_ERRMSG-3-RSRC_ERR: Chassis <num> R0/0: fed: Failed to allocate hardware resource for adj entry
```

驗證由SSDP引起的硬體資源耗盡

所有Catalyst 9000系列交換器都使用特殊的ASIC，以高吞吐量執行大部分封包路由。這些ASIC利用容量有限的不同表和內部資源。由於SSDP客戶端同時充當公共組播組的源和接收器，因此硬體必須使用這些有限的資源在硬體中程式設計路徑以便資料包跟蹤，即使這些資料包由於其他原因從

未到達或被丟棄(TTL 1)。一旦硬體資源耗盡，任何組（無論其與SSDP的關係如何）均無法安裝新的更新或新增。大量未安裝的SSDP更新（狀態突變）也可能會在軟體中排隊，這也會導致非組播流量的硬體更新中斷或失敗，從而影響使用者流量並導致網路中斷。

如果您的網路配置了PIM，並且對於公認的SSDP組地址具有第3層組播狀態，則本文檔才相關。要驗證此條件，請運行命令 "**show ip mroute 239.255.255.250**" （如有必要，請新增vrf語句）。組239.255.255.250特定於SSDP協定。

如果命令輸出包含大量傳出介面和/或具有此特定組的大量唯一源，則表明系統和網路容易受到SSDP導致的中斷的影響。傳出介面和唯一源的數量越多，影響服務的機會就越高。

圖2: 輸出示例 "show ip mroute 239.255.255.250" 命令，在網路上啟用SSDP。

```
<#root>

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
```

除非將SSDP用於特定用途，否則此輸出應為空，或者傳出介面數量較少和/或唯一源數量較少，以防止資源耗盡和可能的服務影響。

如果看到大量組播組，可以使用命令「show platform software object-manager fp active statistics」或「show platform software object-manager fp switch active statistics」來判斷是否已經耗盡硬體資源。

注意：此命令並非特定於由組播流量觸發的資源耗盡，其他問題可能導致這些值不為零。

圖3: 的輸出 "show platform software object-manager fp active statistics"處於問題狀態

<#root>

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

圖3的輸出顯示了交換機資源耗盡的症狀。在正常操作期間，有幾個命令輸出行是不需要的：

- Pending-issue：此值預計為零或接近零。如果這在該命令的多個迭代中仍是一個較大的非零

值，則表明資源已耗盡
- 掛起的確認：此值應為零或接近零。如果這在該命令的多個迭代中仍是一個較大的非零值，則表明資源已耗盡
- Childless-delete-objects：應為零或接近零。不應為10+的值。
- Error-objects：該值應為零或接近零。不應為10+的值。

在存在大量「待決問題」或「待決確認」計數器的狀態下，始終會增加硬體被誤程式設計的風險。硬體程式設計不正確是單播和組播流量的常見中斷來源。

指令 "**show platform hardware fed switch active fwd-asic resource utilization**" or in some models "**show platform hardware fed active fwd-asic resource utilization**" 可用於檢視ASIC上正在使用的有限資源，並確定內部資源是否已耗盡：

圖4: 輸出示例"show platform hardware fed active fwd-asic resource utilization"一個資源快耗盡了。

<#root>

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
```

在圖4中，「RSC_RIL_INDEX」的值顯示有131065個條目在使用中，只有7個是空閒的。此資源被大量唯一SSDP組佔用。雖然不是特定於SSDP，但空閒條目數少和已分配條目數多的資源表示交換機接近容量問題，必須進行調查。

指令 "**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**" 可用於按資源檢視每個ASIC的使用率細分。SSDP耗盡的另一個可能特徵碼是「Used Values」列，用於使「L3組播條目」接近或位於「Max Values」。

圖5: 輸出示例"show platform hardware fed active fwd-asic resource tcam utilization"在正常操作中

<#root>

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
```

防止由SSDP引起的資源耗盡

要停止資源耗盡，必須在建立第一個L3跳和多播狀態之前停止SSDP流量。最快的解決方案是使用在入口上應用於所有配置了PIM並看到此流量的第3層介面的IPv4訪問控制清單(ACL)。使用show ip mroute 239.255.255.250命令進行驗證，然後檢視每個組的「Incoming Interface」。這表示流量的來源是哪個L3介面，並知道可以有多個唯一的來源介面。此配置示例允許SSDP在第2層工作，並允許L2相鄰主機發現PNP服務，但阻止跨第3層邊界轉發客戶端通告，以及阻止在任何組播路由器或交換機上建立L3組播狀態。

設定延伸型ACL:

```
<#root>

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
```

在每個L3介面下設定，請在輸入方向套用ACL:

```
<#root>
```

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

阻止SSDP的備用方法

存在其他方法來限制或完全阻止從SSDP流量建立狀態。因為每個網路不同，並非所有網路都同等有效，而且每個環境都可能具有某些獨特的優勢或劣勢。在撰寫本文時，在SVI攔截流量的路由ACL仍然是最建議、最有效、配置最密集的，以實現減少此流量的狀態和容量的目標，同時仍允許終端客戶端使用此協定發現其本地vlan上的服務。

仔細瞭解每種方法的優點和缺點，以確定它們是否更適合您的環境。

替代方法1：配置PIM RP過濾器以阻止向RP註冊SSDP

此方法對於具有靜態交匯點(RP)對映的環境非常有用，因為在這種對映中，跨大量SVI或第3層介面建立ACL可能是配置密集型操作。

- 此方法的優點是它允許單個配置一次應用於多個第3層介面。
- 此方法的缺點在於，作為第一跳路由器正常狀態建立的一部分，流量仍會被傳送到交換機的CPU。在具有大量直接或間接連線的使用者或大量SSDP流量的環境中，此突發流量仍會與其他合法網路流量競爭用於CPU資源。如果流量保持高位，SSDP流量過多可能會對合法組播流量造成服務影響。

若要實現此方法，請使用以下步驟：

配置拒絕不想要的SSDP流量的ACL:

```
<#root>
Switch(config)#
ip access-list standard 10
Switch(config-std-nacl)#
deny 239.255.255.250 <-- Deny SSDP from registering
Switch(config-std-nacl)#
permit 224.0.0.0 15.255.255.255
```

```
<-- Permit any other group
```

配置作為RP靜態對映的一部分建立的ACL

```
<#root>

Switch#

configure terminal

Switch(config)#

ip pim rp-address 192.168.1.1 10

Switch(config-if)#

end
```

替代方法2：配置Vlan訪問對映(VACL)以拒絕所有SSDP流量

此方法適用於第2層或第3層不需要SSDP的環境，或使用SSDP流量耗盡IGMP監聽或交換器的其他L2多點傳送資源的環境。

- 此方法的優點是易於在單一配置中擴展到大量VLAN。它也是丟棄來自網路的所有SSDP流量的最有效的方式。
- 此方法的缺點是合法使用SSDP發現L2相鄰服務的客戶端現在將無法正常工作。第2層和第3層介面上的所有SSDP流量都會被丟棄，而且不會形成第2層或第3層上的任何狀態。此配置在阻止從本機L3介面上接收的流量建立狀態方面無效。

設定兩個ACL。一個必須僅與SSDP流量匹配，另一個必須是catch-all用於標識所有正常網路流量。

```
<#root>

Switch(config)#

ip access-list extended match_ssdp

Switch(config-ext-nacl)#

permit ip any host 239.255.255.250

Switch(config-ext-nacl)#

exit
Switch(config)#ip access-list extended match_all

Switch(config-ext-nacl)#

permit ip any any
```

使用兩個序列號配置vlan訪問對映。一個用於拒絕SSDP，一個用於允許所有其他流量。將此方法應用於所需的VLAN。

<#root>

Switch#

configure terminal

Switch(config)#

vlan access-map block_ssdp 10

Switch(config-access-map)#

match ip address match_ssdp

Switch(config-access-map)#

action drop

Switch(config-access-map)#

vlan access-map block_ssdp 20

Switch(config-access-map)#

match ip address match_all

Switch(config-access-map)#

action forward

Switch(config-access-map)#

exit

Switch(config)#

vlan filter block_ssdp vlan-list

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。