在Catalyst 9000系列交换机上实施SSDP最佳实践

目录

[简介](#)
[先决条件](#)
[要求](#)
[使用的组件](#)
[了解企业环境中的SSDP风险](#)
[硬件资源耗尽的症状](#)
[验证SSDP导致的硬件资源耗尽](#)
[防止SSDP导致资源耗尽](#)

简介

本文档介绍旨在丢弃或限制Catalyst 9000系列交换机上的简单服务发现协议(SSDP)数据包的最佳实践配置。

先决条件

要求

Cisco 建议您了解以下主题：

- 协议无关组播(PIM)操作
- SSDP如何针对您的环境使用

使用的组件

本文档中的信息基于以下软件和硬件版本：

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

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

了解企业环境中的SSDP风险

通常，笔记本电脑和移动电话等最终用户设备会自动通告其使用SSDP协议的通用即插即用(UPnP)功能。客户端向IP地址239.255.255.250发送组播通告数据包。这些通告通常以生存时间

(TTL)1发送，并且不超出生成组播数据包的主机的本地子网。要接收网络上其他设备的通告，终端还会向239.255.255.250地址发送IGMP成员身份报告，该报告告诉网络，从任何其他组播源发送到此IP地址的组播流量也必须转发到此客户端。

在包含数百或数千个终端的企业环境中，这些终端都同时充当此组的源和感兴趣的接收方，如果不加以限制，此客户端活动很容易使网络设备不堪重负，并且在网络资源耗尽后可能导致网络中断。

这种疲劳主要通过两种方式之一发生：

1. 触发辅助协议故障的硬件资源耗尽
2. SSDP的接口和平台带宽耗尽，用作分布式拒绝服务(DDoS)攻击。

虽然本文未详细讨论，但必须注意，由于SSDP的开放性，攻击者可能会向启用了此服务的一组客户端发送精心编制的数据包，以触发向一个或多个目标主机发送大响应。创建的大量传出接口状态也意味着少量组播流量会显着降低交换机性能容量，因为交换机需要为特定应用集成电路(ASIC)内的每个传出接口制作一个帧副本。 传出接口列出，20个或更多接口存在更大的容量问题和数据包丢失风险。

硬件资源耗尽的症状

当资源耗尽时，Catalyst 9000系列交换机会打印提及“fman_fp_image”或“FMFP”的系统日志。当交换机出现资源耗尽且需要进一步调查时，可以打印这些错误中的某些或全部。

这些是资源耗尽期间出现的一些较常见错误，但不是全面的列表。

图 1：显示的最常见错误示例，这些错误是交换机资源耗尽的证据

```
%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>
```

验证SSDP导致的硬件资源耗尽

所有Catalyst 9000系列交换机都使用特殊ASIC以高吞吐量执行大多数数据包路由。这些ASIC利用容量有限的不同表和内部资源。由于SSDP客户端同时充当公共组播组的源和接收方，因此硬件必须使用这些有限的资源在硬件中编程路径，以便数据包跟随，即使这些数据包从不会由于其他原因(TTL 1)来到或丢弃。 硬件资源耗尽后，不能安装任何组的新更新或添加，无论其与SSDP的关系如何。大量未安装的SSDP更新（状态波动）也可以在软件中排队，这也可能导致非组播流量的硬件更新中断或失败，从而影响用户流量并导致网络中断。

本文档仅在网络配置了PIM且具有已知SSDP组地址的第3层组播状态时相关。要检验此条件，请运行命令 "**show ip mroute 239.255.255.250**" （如有必要，添加vrf语句）。 组239.255.255.250特定于SSDP协议。

如果命令输出包含大量传出接口和/或此特定组具有大量唯一源，则表明系统和网络容易因SSDP而中断。传出接口和唯一源的数量越多，这可能对服务产生影响的机率就越高。

图 2： 输出示例 "show ip mroute 239.255.255.250" 命令，其中SSDP在网络上处于活动状态。

```
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SSDP用于特定目的，否则此输出应为空，或者传出接口数量少，和/或唯一源数量少，以防止资源耗尽和可能的服务影响。

如果看到大量组播组，可使用命令**show platform software object-manager fp active statistics**或**show platform software object-manager fp switch active statistics**来判断硬件资源是否已耗尽。

注意：此命令不特定于组播流量触发的资源耗尽，其他问题可能导致这些值为非零。

图 3： 输出 "show platform software object-manager fp active statistics"问题状态

```
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
```

图3的输出显示了资源耗尽的交换机的症状。有几条命令输出行在正常操作期间不预期出现：

- 待处理问题：预计该值为零或接近零。如果在命令的多次迭代中此值仍为大的非零值，则表示资源耗尽
- 待确认：预计该值为零或接近零。如果在命令的多次迭代中此值仍为大的非零值，则表示资源耗尽
- 无子对象：预计该值为零或接近零。值不应超过10。
- 错误对象：预计该值为零或接近零。值不应超过10。

在存在大量“待处理问题”或“待处理确认”计数器的状态下，始终会增加硬件被错误编程的风险。错误编程的硬件是单播和组播流量的常见故障来源。

命令 "**show platform hardware fed switch active fwd-asic resource utilization**" or in some models "**show platform hardware fed active fwd-asic resource utilization**" 可用于查看ASIC上使用的一些有限资源并确定内部资源是否已耗尽：

图 4： 输出示例"show platform hardware fed active fwd-asic resource utilization"资源几近耗尽。

```
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
```

在图4中，“RSC_RIL_INDEX”的值显示有131065个条目正在使用，只有7个可用。此资源被大量唯一SSDP组消耗。虽然并非特定于SSDP，但具有较少可用条目和大量已分配条目的资源表明交换机接近容量问题，因此必须进行调查。

命令 "**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**" 可用于查看按资源划分的每个ASIC利用率细分。SSDP耗尽的另一个可能签名是“L3组播条目”的“已用值”列，该列接近或位于“最大值”。

图 5： 输出示例"show platform hardware fed active fwd-asic resource tcam utilization"正常运行

```
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SSDP导致资源耗尽

要停止资源耗尽，必须在创建第一个L3跳和组播状态之前停止SSDP流量。最快的解决方案是使用IPv4访问控制列表(ACL)，该列表应用于所有配置了PIM的L3接口的入口，该接口可以看到此流量。使用**show ip mroute 239.255.255.250**命令验证，并查看每个组的“传入接口”。这表示流量源来自哪个L3接口，并且知道可以有多个唯一源接口。此配置示例允许SSDP在第2层工作，并允许第2层相邻主机发现PNP服务，但阻止客户端通告跨第3层边界转发，并防止在任何组播路由器或交换机上创建第3层组播状态。

配置扩展ACL:

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**

在每个L3接口下配置，在入口方向应用ACL:

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