Implement SSDP Best Practices on Catalyst 9000 Series Switches

Contents

Introduction

This document describes best practices to drop or limit Simple Service Discovery Protocol (SSDP) packets on Catalyst 9000 series switches.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- Protocol Independent Multicast (PIM) operation
- How SSDP is used specific to your environment

Components Used

The information in this document is based on these software and hardware versions:

- 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. If your network is live, ensure that you understand the potential impact of any command.

Understand SSDP Risks In Enterprise Environments

In general, end user devices such as laptops and mobile phones automatically advertise their Universal Plug-and-Play (UPnP) capabilities that use the SSDP protocol. Clients send a multicast advertisement packet to the IP address of 239.255.255.250. These advertisements are often sent with a Time to Live (TTL) of 1, and do not go beyond the local subnet of the hosts that generated the multicast packet. To receive the advertisements of other devices on the network, endpoints also send an IGMP Membership Report to the 239.255.255.250 address, which tells the network that multicast traffic sent to this IP address from any other multicast source must also be forwarded to this client.

In enterprise environments that contain hundreds or thousands of endpoints all acting as both a source, and an interested receiver of this group, this client activity can easily overwhelm network devices if left unchecked and can cause outages once network resources have been exhausted.

This exhaustion primarily happens in one of two ways:

1. Hardware resource exhaustion that triggers secondary protocol failures
2. Interface and platform bandwidth exhaustion from SSDP used as a Distributed Denial of Service (DDoS) attack.

While not discussed in detail in this document, it must be noted that because of the open nature of SSDP, it is possible for an attacker to send a crafted packet to a group of clients with this service enabled in order to trigger a large response be sent to one or a group of destination hosts. The large amount of outgoing interface state that is created also means switch performance capacity can be significantly stressed from a small amount of multicast traffic since the switch is required to make one copy of each frame for each outgoing interface within the Application Specific Integrated Circuit (ASIC). Outgoing interface lists that number 20 or more interfaces run a higher risk of capacity problems and packet loss.

Symptoms of Hardware Resource Exhaustion

Catalyst 9000 series switches print syslogs that mention "fman_fp_image" or "FMFP" when resources have been exhausted. Some, or all, of these errors can be printed when the switch has experienced a resource exhaustion and need to be investigated further.

These are some of the more common errors seen during resource exhaustion but is not a comprehensive list.

Figure 1: Sample of the most common errors printed that are evidence of resource exhaustion on a switch

```
%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
```

Verify Hardware Resource Exhaustion Caused by SSDP

All Catalyst 9000 series switches utilize special ASICs to perform the majority of packet routing at high throughput. These ASICs leverage different tables and internal resources which are finite in their capacity. Because SSDP clients act as both sources and receivers for a common multicast group, the hardware must use these limited resources to program a path in hardware for packets to follow, even if those packets never come or get dropped for other reasons (TTL 1). Once hardware resources are exhausted, no new updates or additions for any group, regardless of its relation to SSDP, can be installed. Large numbers of un-installed SSDP updates (state churn) can also queue up in software, this can also cause hardware updates for non-multicast traffic to be interrupted or fail, which impacts user traffic and causes network outages.

This document is only relevant if your network is configured with PIM and has layer-3 multicast state for the well known SSDP group address. To verify this criteria, run the command "**show ip mroute 239.255.255.250**" (add vrf statements if necessary). The group 239.255.255.250 is specific to the SSDP protocol.

If the command output contains a large number of outgoing interfaces and/or has a large number of unique sources for this specific group, that indicates the system and network are vulnerable to outages caused by SSDP. The higher the number of outgoing interfaces and unique sources, the higher the chances that this can become service impacting.

Figure 2: Sample output of "show ip mroute 239.255.255.250" command with SSDP active on the network.

<#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
```

Unless SSDP is used for a specific purpose, this output is expected be empty, or have a low number of outgoing interfaces and/or have a low number of unique sources in order to prevent resource exhaustion and possible service impacts.

If a large number of multicast groups are seen, the command "**show platform software object-manager fp active statistics**" or "**show platform software object-manager fp switch active statistics**" can be used to tell if a hardware resource has been exhausted.

Note: This command is not specific to resource exhaustion triggered by multicast traffic, other issues can cause these values to be non-zero.

Figure 3: Output of "show platform software object-manager fp active statistics"in problem state

```
<#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
```

The output of figure 3 demonstrates symptoms of a switch with resource exhaustion. There are several command output lines which are not expected during normal operation:

- Pending-issue: This is expected to be zero, or close to it. If this remains a large, non-zero value over several iterations of the command, that is a sign of resource exhaustion
- Pending-acknowledgement: This is expected to be zero, or close to it. If this remains a large, non-zero value over several iterations of the command, that is a sign of resource exhaustion
- Childless-delete-objects: This is expected to be zero or close to it. Values of 10+ are not expected.
- Error-objects: This is expected to be zero or close to it. Values of 10+ are not expected.

Consistently in a state where there are large numbers of "pending-issue" or "pending-acknowledgement" counters increases the risk the hardware becomes mis-programmed. Incorrectly programmed hardware is a common source of outages to unicast and multicast traffic.

The command "**show platform hardware fed switch active fwd-asic resource utilization**" or in some models "**show platform hardware fed active fwd-asic resource utilization**" can be used to look at some of the finite resources in use on the ASICs and determine if an internal resource has been exhausted:

Figure 4: Sample output of"show platform hardware fed active fwd-asic resource utilization"with one resource near exhaustion.

<#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
<snip>
```

In figure 4 the value for "RSC_RIL_INDEX" shows there are 131065 entries in use, and only 7 are free. This resource is consumed by large numbers of unique SSDP groups. While not specific to SSDP, resources that have a low number of free entries and high number allocated entries are signs the switch is near a capacity issue, and must be investigated.

The command "**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**" can be used to look at a per-ASIC breakdown of utilization by resource. Another possible signature from SSDP exhaustion is the "Used Values" column for "L3 Multicast entries" to close to or at the "Max Values".

Figure 5: Sample output of"show platform hardware fed active fwd-asic resource tcam utilization"in normal operation

```
<#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
```

Prevent Resource Exhaustion Caused by SSDP

To stop resource exhaustion, the SSDP traffic must be stopped prior to the first L3 hop and multicast state creation. The quickest solution is to use an IPv4 Access Control List (ACL) applied on ingress to all L3 interfaces configured with PIM that sees this traffic. Verify with the "**show ip mroute 239.255.255.250**" command and look at the "Incoming Interface" for each group. This indicates which L3 interface the source of the traffic is sourced from and be aware there can be more than one unique source interface. This configuration example allows SSDP to work at layer 2 and allows L2-adjacent hosts to discover PNP services, but prevents client advertisements to be forwarded across L3 boundaries, and prevents L3 multicast state creation on any multicast router or switch.

Configure an extended 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
```

Configure under each L3 interface, apply the ACL in the ingress direction:

<#root>

```
Switch#

configure terminal

Switch(config)#
```

```
interface vlan100

Switch(config-if)#

ip access-group BLOCK_SSDP in

Switch(config-if)#

end
```

Alternative Methods Blocking SSDP

Other methods exist to limit or completely prevent the creation of state from SSDP traffic. Because each network is different, not all are equally as effective, and can come with certain advantages or disadvantages unique to each environment. At the time of this writing, a routed ACL blocking traffic at the SVI remains the most recommended, most effective, and least configuration intensive to accomplish the goal of reducing state and volume of this traffic, while still allowing end clients to use this protocol to discover services on their local vlan.

Carefully understand the advantages and disadvantages of each of the methods to determine if one can be a better fit for your environment.

Alternative Method 1: Configure PIM RP Filter to Prevent SSDP Registration With the RP

This method is useful for environments with a static Rendezvous Point (RP) mapping where the creation of an ACL across a large number of SVIs or L3 interfaces can be configuration intensive.

- The advantage of this method is that it allows a single configuration to apply across multiple L3 interfaces at once.
- The disadvantage of this method is that traffic is still punted to the CPU of the switch as part of normal state creation for a first-hop router. In environments with large amounts of directly or indirectly connected users or large amounts of SSDP traffic, this punted traffic still competes with other legitimate network traffic for CPU resources. Excessive volumes of SSDP traffic can cause a service impact to legitimate multicast traffic if the volume of traffic remains high.

To implement this method, use these steps:

Configure an ACL denying undesirable SSDP traffic:

```
<#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
```

Configure the ACL you created as part of the RP static mapping

<#root>

```
Switch#

configure terminal

Switch(config)#

ip pim rp-address 192.168.1.1 10

Switch(config-if)#

end
```

Alternative Method 2: Configure Vlan Access-Maps (VACL) to Deny All SSDP Traffic

This method is useful for environments where SSDP is not needed at L2 or L3, or in environments where the volume of SSDP traffic exhausts IGMP snooping or other L2 multicast resources of the switch.

- The advantage of this method is that it is easy to scale to a large number of vlans in a single configuration. It is also the most effective form of discarding all SSDP traffic from the network.
- The disadvantage of this method is that clients who legitimately use SSDP to discover L2 adjacent services will now fail to function properly. All SSDP traffic on both L2 and L3 interfaces will be discarded and no state at L2 or L3 will be formed. This configuration is not effective at blocking state creation from traffic received on native L3 interfaces.

Configure two ACLs. One must match only SSDP traffic, and one must be a catch-all used to identify all normal network traffic.

<#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
```

Configure a vlan access map with two sequence numbers. One to deny SSDP, one to permit all other traffic. Apply this to the desired vlans.

<#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 <vlan(s)|all>
```