

某局点S5800-32C Ingress LSP和ECMP 资源超阈值问题

LSP

产品特性

折锐鹏

2020-04-14 发表

组网及说明

不涉及

问题描述

现场MPLS L3VPN组网，近期PE设备5800出现Ingress LSP超过阈值的告警，此外，查看local logbuffer信息，ECMP也超资源阈值了。

log日志中有大量Ingress LSP超过阈值的日志。

LSPM/5/LSPM_LSP_LIMIT: Ingress LSP limit reached.

通过命令查看发现ingress达到1024条。

```
<TJZW-W-JR-GaoFa-5800>dis mpls lsp statistics
Lsp Type      Total      Ingress    Transit    Egress
STATIC LSP     0           0           0           0
STATIC CRLSP   0           0           0           0
LDP LSP        1182        1024        155         3
CRLDP CRLSP    0           0           0           0
RSVP CRLSP     0           0           0           0
BGP LSP        10          0           0          10
ASBR LSP       0           0           0           0
BGP IPV6 LSP   0           0           0           0
-----
LSP            1192        1024        155         13
CRLSP          0           0           0           0
<TJZW-W-JR-GaoFa-5800>
```

此外，在local lobuffer里，有如下ECMP资源不足的告警。

Mar 30 2020 13:09:54:0396:

LINE:1967, File drv_ip4_shim.c-TASK:FIB-FUNC::

call DRV_L3UC_ECMP_Malloc error:No resource to add ECMP entry!

过程分析

Ingress LSP超阈值（规格为1024）是因为PE 5800设备上配置了lsp-trigger all命令，这个时候公网侧不仅32位掩码的主机路由能够触发LDP建立LSP，网段路由也可以建立LSP。从诊断看，建了很多的网段EFC的LSP。一般情况下公网侧只要保证两个PE的loopback地址之间建立LSP，从而可以将报文按照LSP转发到对端PE侧即可。因为现场大量的网段路由也建立LSP，从而占用了很多ingress lsp资源。查看display mpls lsp，可以发现有很多网段路由的LSP。

=====display mpls lsp=====

LSP Information: LDP LSP

FEC	In/Out Label	In/Out IF	Vrf Name
100.65.90.4/30	NULL/39135	-/Vlan3000	
100.65.100.2/32	NULL/39136	-/Vlan3000	
192.168.49.0/24	NULL/1025	-/Vlan3000	
100.65.70.57/32	NULL/3574	-/Vlan3000	
100.65.70.21/32	NULL/2283	-/Vlan3000	
100.65.70.20/32	NULL/4002	-/Vlan3000	
100.65.8.84/30	NULL/40214	-/Vlan3000	
100.65.8.28/30	NULL/40236	-/Vlan3000	
100.65.8.164/30	NULL/39497	-/Vlan3000	
100.65.6.0/30	NULL/39861	-/Vlan3000	
100.65.51.80/30	NULL/38980	-/Vlan3000	
100.65.0.128/32	NULL/39862	-/Vlan3000	
100.65.149.120/29	NULL/39627	-/Vlan3000	
100.65.149.32/29	NULL/40220	-/Vlan3000	
100.65.2.136/32	NULL/39068	-/Vlan3000	
100.65.149.160/29	NULL/39587	-/Vlan3000	
100.65.1.34/32	NULL/39538	-/Vlan3000	
100.65.1.9/32	NULL/40274	-/Vlan3000	

#

mpls

lsp-trigger all

#

lsp-trigger

【命令】

lsp-trigger [vpn-instance vpn-instance-name] { all | ip-prefix prefix-name }

undo lsp-trigger [vpn-instance vpn-instance-name] { all | ip-prefix prefix-name }

【视图】

MPLS视图

【缺省级别】

2: 系统级

【参数】

vpn-instance *vpn-instance-name*: 配置指定VPN的LSP触发策略。其中，*vpn-instance-name*表示MPLS L3VPN的VPN实例名称，为1~31个字符的字符串，区分大小写。

all: 所有路由项都会触发LDP建立LSP。

ip-prefix *prefix-name*: 利用IP地址前缀列表对路由项进行过滤，被IP地址前缀列表拒绝的静态路由和IGP路由项不能触发建立LSP。*prefix-name*表示IP地址前缀列表名，为1~19个字符的字符串。

【描述】

lsp-trigger命令用来配置LSP的触发策略。**undo lsp-trigger**命令用来恢复缺省情况。

缺省情况下，只有32位掩码的主机路由能够触发LDP建立LSP。

需要注意的是：

- IP地址前缀列表只对静态路由和IGP路由有效。
- LSP的建立需要LSR上有精确匹配的路由项，如果使用32位掩码的Loopback接口，则必须有精确匹配的主机路由才能触发LSP的建立。
- 如果指定**vpn-instance** *vpn-instance-name*参数，则配置指定VPN的LSP触发策略；如果未指定**vpn-instance** *vpn-instance-name*参数，则配置公网路由的LSP触发策略。
- 有关IP地址前缀列表的介绍请参见“三层技术-IP路由”

关于ECMP超资源，是因为现场很多路由采用了ECMP。因为V5的5800设备没有ECMP复用技术，所以针对不同的目的地址，如果有多个下一跳，则都会占用ECMP资源的。新的一些V7交换机产品ECMP是可以复用的，同样的下一跳复用一条资源。

通过如下命令可以看到ecmp已经使用了255条了。

=====l3 multipath show=====

Multipath Egress Object 200000

Ecmp Cnt: 2

Interfaces: 101157 100438

.....

.....

Multipath Egress Object 200254

Ecmp Cnt: 2

Interfaces: 100626 100056

解决方法

针对Ingress LSP超阈值的问题，现场取消掉lsp-trigger命令后，ingress LSP的数目下降到了465。

```
[TJZW-W-3R-Gaofa-5800]dis mpls lsp statistics
Lsp Type      Total      Ingress    Transit    Egress
STATIC LSP    0           0           0           0
STATIC CRLSP  0           0           0           0
LDP LSP       1182        1024        155         3
CRLDP CRLSP   0           0           0           0
RSVP CRLSP    0           0           0           0
BGP LSP       10          0           0           10
ASBR LSP      0           0           0           0
BGP IPv6 LSP  0           0           0           0
-----
LSP           1192        1024        155         13
CRLSP         0           0           0           0
[TJZW-W-3R-Gaofa-5800]mpls
[TJZW-W-3R-Gaofa-5800-mpls]dis th
#
mpls
lsp-trigger all
return
[TJZW-W-3R-Gaofa-5800-mpls]undo lsp
[TJZW-W-3R-Gaofa-5800-mpls]undo lsp-trigger all
[TJZW-W-3R-Gaofa-5800-mpls]
[TJZW-W-3R-Gaofa-5800-mpls]dis mpls lsp statistics
Lsp Type      Total      Ingress    Transit    Egress
STATIC LSP    0           0           0           0
STATIC CRLSP  0           0           0           0
LDP LSP       535         465         69          1
CRLDP CRLSP   0           0           0           0
RSVP CRLSP    0           0           0           0
BGP LSP       10          0           0           10
ASBR LSP      0           0           0           0
BGP IPv6 LSP  0           0           0           0
-----
LSP           545         465         69          11
CRLSP         0           0           0           0
```

针对ECMP超资源，目前只能通过缩减ECMP路由的使用来解决。