



S6800 聚合出接口hash不均

二层链路聚合

软件问题

李鹏飞

2022-07-14 发表

组网及说明

过路报文为vxlan报文，源目IP固定，源端口离散

问题描述

BAGG1口有如下成员口

Aggregate Interface: Bridge-Aggregation1

Aggregation Mode: Dynamic

Loadsharing Type: Shar

Management VLANs: None

System ID: 0x8000, dcda-8038-0b8c

Local:

Port	Status	Priority	Index	Oper-Key	Flag
XGE1/0/1	S	32768	5	3	{ABCDEF}
XGE1/0/2	S	32768	6	3	{ABCDEF}
XGE1/0/3	S	32768	7	3	{ABCDEF}
XGE1/0/4	S	32768	8	3	{ABCDEF}
XGE1/0/5	S	32768	9	3	{ABCDEF}
XGE1/0/6	S	32768	10	3	{ABCDEF}
XGE1/0/7	S	32768	11	3	{ABCDEF}
XGE1/0/8	S	32768	12	3	{ABCDEF}
XGE1/0/9	S	32768	13	3	{ABCDEF}
XGE1/0/10	S	32768	14	3	{ABCDEF}
XGE1/0/11	S	32768	15	3	{ABCDEF}
XGE1/0/12	S	32768	16	3	{ABCDEF}
XGE1/0/13	S	32768	17	3	{ABCDEF}
XGE1/0/14	S	32768	18	3	{ABCDEF}
XGE1/0/15	S	32768	19	3	{ABCDEF}
XGE1/0/16	S	32768	20	3	{ABCDEF}

出接口流量严重不均

Interface	Usage (%)	Total (pps)	Broadcast (pps)	Multicast (pps)
BAGG1	5	1267425	--	--
XGE1/0/1	0	8138	--	--
XGE1/0/2	0	162	--	--
XGE1/0/3	1	10974	--	--
XGE1/0/4	0	250	--	--
XGE1/0/5	0	1710	--	--
XGE1/0/6	0	449	--	--
XGE1/0/7	75	1154399	--	--
XGE1/0/8	0	270	--	--
XGE1/0/9	5	72886	--	--
XGE1/0/10	0	766	--	--
XGE1/0/11	0	340	--	--
XGE1/0/12	0	232	--	--
XGE1/0/13	1	15438	--	--
XGE1/0/14	0	765	--	--
XGE1/0/15	0	353	--	--
XGE1/0/16	0	293	--	--

过程分析

1. 查看现场vxlan报文hash计算的hash因子的寄存器。

现场寄存器

```
bcm s 1 c 0 g/RTAG7_IPV4_TCP_UDP_HASH_FIELD_BMAP_2
RTAG7_IPV4_TCP_UDP_HASH_FIELD_BMAP_2.ipipe0[1][0x16001c00]=0x1ef8f70
: <IPV4_TCP_UDP_FIELD_BITMAP_B=0xf7c,IPV4_TCP_UDP_FIELD_BITMAP_A=0xf00>
```

实验室情况:

```
[H3C-probe]bcm s 1 c 0 g/RTAG7_IPV4_TCP_UDP_HASH_FIELD_BMAP_2
RTAG7_IPV4_TCP_UDP_HASH_FIELD_BMAP_2.ipipe0[1][0x16001c00]=0x1ef8f70
: <IPV4_TCP_UDP_FIELD_BITMAP_B=0xf7c,IPV4_TCP_UDP_FIELD_BITMAP_A=0xf70>
```

源端口号对应的bit位是0x40(二进制0x01000000)，现场hash的寄存器0xf00,不包含0x40，也就是源端口号没有参与hash因子计算

实验室用修改寄存器命令修改寄存器和现场一样现场，可以复现现场现象。

用 bcm s 1 c 0

```
m/RTAG7_IPV4_TCP_UDP_HASH_FIELD_BMAP_2/IPv4_TCP_UDP_FIELD_BITMAP_A=0xf70 命令将寄存器值改对，可以使问题恢复。
```

进一步排查现场的寄存器为什么是0xf00。发现在R2612P02H11 至R2612P02H21 之间的补丁上，合入了聚合对称HASH的功能，此功能会将寄存器值修改为0xf00，会使默认HASH不包含L4源目的端口号。

1.1 R2612P02H11版本操作方式变更

1. [201901210102]增加IP报文的对称Hash功能

变更前：对IPv4的非已知单播流量无法做到对称Hash。

变更后：对IPv4的非已知单播流量可以做到对称Hash。

解决方法

后续H2后续H23及更高的补丁上取消的这个修改，不会修改默认寄存器值。但是如果设备已经安装过H11-H21的补丁，再安装H23及更高补丁，不会再把这个寄存器值主动修改回来。也就是说现场安装了H19补丁，后续再安装H23以上的补丁，仍然存在此次问题。

现场设备出问题设备可以使用如下命令修改恢复。

```
bcm slot x chip 0
```

```
m/RTAG7_IPV4_TCP_UDP_HASH_FIELD_BMAP_2/IPV4_TCP_UDP_FIELD_BITMAP_A=0xf70
```

