

知 某局点S6520X 分片报文不通问题

二层转发

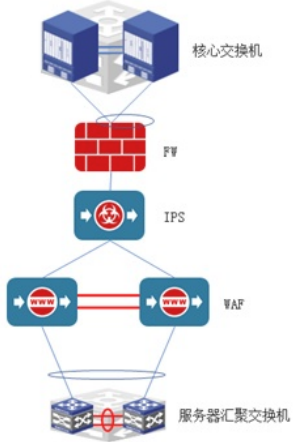
转发不通

苏亚东

2022-06-20 发表

组网及说明

测试终端192.168.212.228



测试服务器192.168.160.168

问题描述

现场组网大致如下，核心交换机12518和S6520X之间经过透明FW和WAF聚合，由于waf侧两台设备都是单机的，如果没有会话就会将报文丢弃。因此就需要S6520X将同一条流发给同一台waf进行转发。目前现场测试发现，192.168.212.228去ping测试192.168.160.168时，小包通信正常，但是大包会出现问题。

过程分析

(1) 故障时在wa侧进行抓包，发现两台wa都能从S6520X收到同一条流的流量。

Frame	Source IP	Destination IP
11	192.168.160.168	192.168.212.228
12	192.168.212.228	192.168.160.168
13	192.168.212.228	192.168.160.168
14	192.168.160.168	192.168.212.228
15	192.168.160.168	192.168.212.228
16	192.168.160.168	192.168.212.228
17	192.168.160.168	192.168.212.228
18	192.168.160.168	192.168.212.228
19	192.168.160.168	192.168.212.228
20	192.168.160.168	192.168.212.228
21	192.168.160.168	192.168.212.228
22	192.168.160.168	192.168.212.228
23	192.168.160.168	192.168.212.228
24	192.168.160.168	192.168.212.228
25	192.168.160.168	192.168.212.228
26	192.168.160.168	192.168.212.228
27	192.168.160.168	192.168.212.228
28	192.168.160.168	192.168.212.228
29	192.168.160.168	192.168.212.228
30	192.168.160.168	192.168.212.228
31	192.168.160.168	192.168.212.228
32	192.168.160.168	192.168.212.228
33	192.168.160.168	192.168.212.228
34	192.168.160.168	192.168.212.228
35	192.168.160.168	192.168.212.228
36	192.168.160.168	192.168.212.228
37	192.168.160.168	192.168.212.228
38	192.168.160.168	192.168.212.228
39	192.168.160.168	192.168.212.228
40	192.168.160.168	192.168.212.228
41	192.168.160.168	192.168.212.228
42	192.168.160.168	192.168.212.228
43	192.168.160.168	192.168.212.228
44	192.168.160.168	192.168.212.228
45	192.168.160.168	192.168.212.228
46	192.168.160.168	192.168.212.228
47	192.168.160.168	192.168.212.228
48	192.168.160.168	192.168.212.228
49	192.168.160.168	192.168.212.228
50	192.168.160.168	192.168.212.228
51	192.168.160.168	192.168.212.228
52	192.168.160.168	192.168.212.228
53	192.168.160.168	192.168.212.228
54	192.168.160.168	192.168.212.228
55	192.168.160.168	192.168.212.228
56	192.168.160.168	192.168.212.228
57	192.168.160.168	192.168.212.228
58	192.168.160.168	192.168.212.228
59	192.168.160.168	192.168.212.228
60	192.168.160.168	192.168.212.228
61	192.168.160.168	192.168.212.228
62	192.168.160.168	192.168.212.228
63	192.168.160.168	192.168.212.228
64	192.168.160.168	192.168.212.228
65	192.168.160.168	192.168.212.228
66	192.168.160.168	192.168.212.228
67	192.168.160.168	192.168.212.228
68	192.168.160.168	192.168.212.228
69	192.168.160.168	192.168.212.228
70	192.168.160.168	192.168.212.228
71	192.168.160.168	192.168.212.228
72	192.168.160.168	192.168.212.228
73	192.168.160.168	192.168.212.228
74	192.168.160.168	192.168.212.228
75	192.168.160.168	192.168.212.228
76	192.168.160.168	192.168.212.228
77	192.168.160.168	192.168.212.228
78	192.168.160.168	192.168.212.228
79	192.168.160.168	192.168.212.228
80	192.168.160.168	192.168.212.228
81	192.168.160.168	192.168.212.228
82	192.168.160.168	192.168.212.228
83	192.168.160.168	192.168.212.228
84	192.168.160.168	192.168.212.228
85	192.168.160.168	192.168.212.228
86	192.168.160.168	192.168.212.228
87	192.168.160.168	192.168.212.228
88	192.168.160.168	192.168.212.228
89	192.168.160.168	192.168.212.228
90	192.168.160.168	192.168.212.228
91	192.168.160.168	192.168.212.228
92	192.168.160.168	192.168.212.228
93	192.168.160.168	192.168.212.228
94	192.168.160.168	192.168.212.228
95	192.168.160.168	192.168.212.228
96	192.168.160.168	192.168.212.228
97	192.168.160.168	192.168.212.228
98	192.168.160.168	192.168.212.228
99	192.168.160.168	192.168.212.228
100	192.168.160.168	192.168.212.228

(2) 进一步确认设备上相关流量转发情况，发现对应流量在交换机上只进行二层转发。并且，该流量在S6520X上的上、下行口都进行了跨slot的聚合。

下行口：

Aggregate Interface: Bridge-Aggregation1

Aggregation Mode: Static

Loadsharing Type: Shar

Management VLANs: None

Port Status Priority Oper-Key

XGE1/0/3(R) S 32768 1

XGE2/0/3 S 32768 1

(3) 根据抓包信息看到具体流量的五元组等hash因子均相同，如果流量都从同一个slot进来，不太可能又跨slot转发出去，此时怀疑现场服务器两个口都会发送流量，服务器将报文分片后轮流从两个口发往交换机的不同slot，而交换机侧默认聚合口都是本地优转的，即流量从哪一个slot进来，只有该slot有对应出接口，就会从本slot的出接口发出。

(4) 根据该理论，由于服务器侧无法进行相关信息查看，在交换机侧关闭本地优转测试，故障消除。

undo link-aggregation load-sharing mode local-first

由此可以确认，服务器侧两个口都会发包，导致交换机正常本地优转发给远端时，分片报文被中间的wa设备检查不通过导致丢包。

解决方法

- (1) 在交换机侧关闭本地优转：
undo link-aggregation load-sharing mode local-first
- (2) 调整服务器发包模式，只从一个口发出；
- (3) 调整waf的策略，交换机正常聚合口转发不需要考虑流量路径。

