某局点突然内存告警并且OSPF无法建立问题

OSPF 张文宁 2022-01-16 发表

组网及说明

无

问题描述

这台S12508设备下挂业务突然出现异常，同时登录该设备发现日志打印ospf邻居关系down和lowmem内存占用超阈值的日志告警，配置静态路由后业务得以恢复。

D-01 OSPF/5/OSPF_NBR_CHG: OSPF 200 Neighbor 10.87.21.41(Route-Aggregation1) from FULL to DOWN.

Dec 28 07:04:40:018 2021 DC1-SOC-4A10-S12508-D-01 OSPF/5/OSPF_NBR_CHG: OSPF 200 Neighbor 10.87.22.41(Route-Aggregation2) from FULL to DOWN.

Dec 28 07:04:40:024 2021 DC1-SOC-4A10-S12508-D-01 DIAG/4/MEM_ALERT:

system memory info: total used free shared buffers cached

Mem: 3491784 1233240 2258544 0 32 143944

-/+ buffers/cache: 1089264 2402520

Swap: 0 0 0

Lowmem: 1005512 956400 49112

Dec 28 07:04:41:467 2021 DC1-SOC-4A10-S12508-D-01 DIAG/1/MEM_EXCEED_THRESHOLD: Memory critical threshold has been exceeded.

过程分析

从日志分析，设备在12月28日07:04，lowmem内存抵达三级门限告警，触发内存保护从而ospf邻居被down掉。从内存快速降低的情况看，可以确认是lowmem内存泄漏导致。

%Dec 28 07:04:40:024 2021 DC1-SOC-4A10-S12508-D-01 DIAG/4/MEM_ALERT:
system memory info:

	total	used	free	shared	buffers	cached
Mem:	3491784	1233240	2258544	0	32	143944
-/+ buffers/cache:		1089264	2402520			
Swap:	0	0	0			
Lowmem:	1005512	956400	49112			

%Dec 28 07:04:41:467 2021 DC1-SOC-4A10-S12508-D-01 DIAG/1/MEM_EXCEED_THRESHOLD:
Memory critical threshold has been exceeded.

%Dec 28 07:04:40:016 2021 DC1-SOC-4A10-S12508-D-01 OSPF/5/OSPF_NBR_CHG: -Chassis=1-Slot=0; OSPF 200 Neighbor 10.87.21.41(Route-Aggregation1) from FULL to DOWN.
%Dec 28 07:04:40:018 2021 DC1-SOC-4A10-S12508-D-01 OSPF/5/OSPF_NBR_CHG: -Chassis=1-Slot=0; OSPF 200 Neighbor 10.87.22.41(Route-Aggregation2) from FULL to DOWN.

为进一步确认lowmem内存泄漏原因，登录设备多次查看具体模块内存占用情况如下，发现其中0x208模块（聚合模块）内存占用明显异常，并且不断在增长，可以确认该模块存在lowmem内存泄漏。

[DC1-SOC-4A10-S12508-D-01-probe]display system internal kernel memory pool tag chassis 1 slot 1

Tag	MemoryPool	Actives	Bytes
0x11a0000	6	8	1069424
0x1140000	10	803	1474920
0x10b0000	17	4789	1395204
0x2120000	1	1	8
0x23c0000	1	1	8
0x2230000	2	3	16464
0x2080000	6	8353	4221880
0x2300000	3	3	49200

[DC1-SOC-4A10-S12508-D-01-probe] display system internal kernel memory pool tag chassis 1 slot 1

Tag	MemoryPool	Actives	Bytes
0x11a0000	6	8	1069424
0x1140000	10	803	1474920
0x10b0000	16	4682	1349608
0x2120000	1	1	8
0x23c0000	1	1	8
0x2230000	2	3	16464
0x2080000	6	9097	4602328
0x2300000	3	3	49200

[DC1-SOC-4A10-S12508-D-01-probe] display system internal kernel memory pool tag chassis 1 slot 1

Tag	MemoryPool	Actives	Bytes
0x11a0000	6	8	1069424
0x1140000	10	803	1474920
0x10b0000	15	4680	1347564
0x2120000	1	1	8
0x23c0000	1	1	8
0x2230000	2	3	16464
0x2080000	6	9350	4732344
0x2300000	3	3	49200

通过走读代码和分析，确认网管通过snmp读取设备聚合口的两个table下的节点（hh3cIfHCFlowStatTable和hh3cIfFlowStatTable）时，设备响应后，会出现lowmem内存泄露的问题。

lowmem内存泄漏的原因：MIB下发请求获取统计信息，聚合模块处理该请求时，会下发驱动获取聚合成员口相关统计信息，然后累加作为聚合口统计信息；在处理函数内部，会临时申请一块内存用来存储聚合口统计，处理完毕后，会将该数据拷贝回MIB请求的缓冲区中，由于申请和拷

贝不在同一层函数调用中，遗漏释放申请的临时内存，导致内存泄露。

S12508设备在lowmem内存泄漏时不会单独发送trap告警，导致网管无法提前监控到lowmem内存泄漏的告警。

解决方法

1、网管取消通过snmp读取S12500系列交换机的hh3cIfHCFlowStatTable和hh3cIfFlowStatTable两个table下的节点，规避lowmem内存泄漏问题。

2、在设备上配置oid访问控制，拒绝hh3cIfHCFlowStatTable和hh3cIfFlowStatTable两个table下的节点被访问，规避lowmem内存泄漏问题。

3、将S12500系列交换机升级到最新的R7377P04版本，解决lowmem内存泄漏问题。

hh3cIfHCFlowStatTable

OID of this table is: 1.3.6.1.4.1.25506.2.40.2.1.2.3

This table is only supported by physical interfaces and serial interfaces created by controller.

Name	Access	PDS	Description
hh3cIfStatFlowHCInBits (1.3.6.1.4.1.25506.2.40.2.1.2.3.1.1)	read-only	No	As per MIB
hh3cIfStatFlowHCOutBits (1.3.6.1.4.1.25506.2.40.2.1.2.3.1.2)	read-only	No	As per MIB
hh3cIfStatFlowHCInPkts (1.3.6.1.4.1.25506.2.40.2.1.2.3.1.3)	read-only	No	As per MIB
hh3cIfStatFlowHCOutPkts (1.3.6.1.4.1.25506.2.40.2.1.2.3.1.4)	read-only	No	As per MIB
hh3cIfStatFlowHCInBytes (1.3.6.1.4.1.25506.2.40.2.1.2.3.1.5)	read-only	No	As per MIB
hh3cIfStatFlowHCOutBytes (1.3.6.1.4.1.25506.2.40.2.1.2.3.1.6)	read-only	No	As per MIB

hh3cIfFlowStatTable

OID of this table is: 1.3.6.1.4.1.25506.2.40.2.1.2.1

This table is only supported by physical interfaces and serial interfaces created by controller.

Name	Access	PDS	Description
hh3cIfStatFlowInterval (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.1)	read-write	Current	Only support read operation.
hh3cIfStatFlowInBits (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.2)	read-only	No	As per MIB
hh3cIfStatFlowOutBits (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.3)	read-only	No	As per MIB
hh3cIfStatFlowInPkts (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.4)	read-only	No	As per MIB
hh3cIfStatFlowOutPkts (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.5)	read-only	No	As per MIB
hh3cIfStatFlowInBytes (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.6)	read-only	No	As per MIB
hh3cIfStatFlowOutBytes (1.3.6.1.4.1.25506.2.40.2.1.2.1.1.7)	read-only	No	As per MIB

实验室搭建环境模拟，S12508设备使用现网版本R7376可以复现lowmem内存泄漏问题。但升级到R7377P01及以后版本不会复现该问题。进一步走读代码和分析，发现在R7377P01版本中对获取聚合口统计计数相关流程进行过优化调整，已不存在该问题。

R7377P01以前的版本统计获取数据的方式不符合部分场景实现，所以在R7377P01及以后的版本对获取聚合口统计计数相关流程做了优化，不会再申请临时内存，而使用临时变量，因此不会再涉及到lowmem内存泄漏的问题。

综上，设备现网版本代码存在首次发现的漏洞，当网管通过snmp读取hh3cIfHCFlowStatTable和hh3cIfFlowStatTable两个table下的节点时，引起lowmem内存泄漏，可能会导致设备超过三级门限，触发内存保护，使得ospf邻居down，进而影响业务。