管理EOR板卡ND表项死锁问题分析

IPv6 黄荣德 2024-06-30 发表

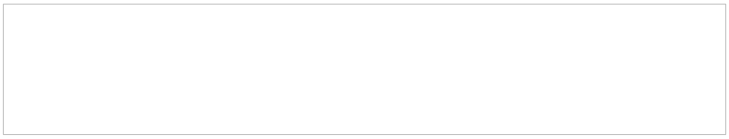
问题描述

一、问题原因

根据前期现网流量抓包分析：

从抓包看，我们设备发送的单播NS探测报文，3都不回应，组播NS网元会回应。为什么不回应，这个得网元侧确认下

应该是这个原因导致EOR设备的ND表项一直在反复的绑定，解绑定



根据复位日志分析：kND队列任务出现死循环，导致设备重起

%Jan 9 10:26:04:747 2023 NFV-D-HDBJN-03A-2402-0E24-M-EOR-01 DRVPLAT/4/DrvDebug: -Slot=8;
Task: CPU 5 is occupied by process kND/1 for more than 15 seconds.

%Jan 9 10:26:05:157 2023 NFV-D-HDBJN-03A-2402-0E24-M-EOR-01 DRNI/6/DRNI_KEEPALIVELINK_DOWN: Keepalive link went down because the peer keepalive timeout timer expired. Please check the keepalive packet transmission and reception status at the two ends.
%Jan 9 10:26:13:227 2023 NFV-D-HDBJN-03A-2402-0E24-M-EOR-01 DIAG/0/DIAG_KDBG: -Slot=8; Deadloop once occurred on slot 8 cpu 0.
%Jan 9 10:26:13:041 2023 NFV-D-HDBJN-03A-2402-0E24-M-EOR-01 DEV/2/BOARD_STATE_FAULT: Board state changed to Fault on slot 8, type is LSXM1TGS48HB0.

二、重起的调用栈：

分析是死在了spin_lock_irqsave锁上

<3>[5131059.989893] 5:BUG: soft lockup - CPU#5 stuck for 23s! [kND/1:446]
<4>[5131059.989951] Modules linked in: ksplice_lpu_xlp_1649245709_system_old ksplice_lpu_xlp_1649245709_system_new ksplice_lpu_xlp_1649245709 system addon
<4>[5131059.990065] Cpu 5
<4>[5131059.990085] $ 0 : 0000000000000000 0000000000000018 ffffffffc0004001 0000000000000001
<4>[5131059.990159] $ 4 : c00000019f289f70 0000000000000100 0000000000000000 0000000000000001
<4>[5131059.990233] $ 8 : c0000001a2871280 000000000000005d 0000000000000000 0000000000000001
<4>[5131059.990309] $12 : 0000000000000030 ffffffff802057c4 0000000000000000 c000000046ed0000
<4>[5131059.990380] $16 : c00000019f289eb0 c0000001a6980480 ffffffffcf35af28 0000000000030fdf
<4>[5131059.990460] $20 : 0000000000000000 ffffffffcf35af28 0000000000000000 c00000019f289f70
<4>[5131059.990532] $24 : 0000000000000000 ffffffff802cb7f0
<4>[5131059.990600] $28 : c00000004acc0000 c00000004accfdf0 0000000000000002 ffffffff804b0294
<4>[5131059.990677] Hi : 0000000000000000
<4>[5131059.990707] Lo : 0000000000000000
<4>[5131059.990739] epc : ffffffff804b0954 _spin_lock_irqsave+0xe4/0x200 Tainted: P
<4>[5131059.990826] ra : ffffffff804b0294 _write_lock_bh+0x14/0x50
<4>[5131059.990878] Status: 5400ffe3 KX SX UX KERNEL EXL IE
<4>[5131059.990939] Cause : 40808000
<4>[5131059.990970] PrId : 000c1104 (XLP308 Rev C0)
<4>[5131059.991001] Modules linked in: ksplice_lpu_xlp_1649245709_system_old ksplice_lpu_xlp_1649245709_system_new ksplice_lpu_xlp_1649245709 system addon
<4>[5131059.991115] Process kND/1 (pid: 446, threadinfo=c00000004acc0000, task=c00000004a17c9f8, tls=0000000000000000)
<4>[5131059.991181] Stack :

<4>[5131059.991203] c00000019f289f70 0000000000000000
<4>[5131059.991247] ffffffffcf35af28 ffffffffc7556f00
<4>[5131059.991291] 0000000000000000 c00000019ec78480
<4>[5131059.991340] c00000019ec78480 ffffffffc75609b0
<4>[5131059.991386] c00000019f289ef0 ffffffff80252618
<4>[5131059.991432] c00000019f289ef0 ffffffffc75608f0
<4>[5131059.991477] ffffffff804ad500 ffffffffc7556df0
<4>[5131059.991526] ffffffff80319100 ffffffff804ad500
<4>[5131059.991573] ffffffffc72b4a70 ffffffffe3e40000
<4>[5131059.991620] c0000000c57f56b0 ffffffffc7604a18
<4>[5131059.991664] c00000019f289cc0 ffffffff804ad500
<4>[5131059.991713] c0000000c57f5680 ffffffff80266750
<4>[5131059.991759] 0000000000000001 ffffffffc7609988
<4>[5131059.991804] c00000019f289cc0 0000000000000000
<4>[5131059.991849] 0000000000000000 ffffffffc72b4a70
<4>[5131059.991899] ffffffffe3e40000 ffffffffc75f7e80
<4>[5131059.991945] c00000019f289cc0 0000000000000000
<4>[5131059.991991] 0000000000000003 c000000002b24640
<4>[5131059.992034] c00000004aaffcb0 0000000000000000
<4>[5131059.992077] ffffffffc75f7d80 0000000000000000
<4>[5131059.992124] 0000000000000000 0000000000000000
<4>[5131059.992168] 0000000000000000 0000000000000000
<4>[5131059.992218] 0000000000000000 ffffffff80266ab0
<4>[5131059.992264] 0000000000000000 0000000000000000
<4>[5131059.992311] 0000000000000000 c00000004accff78
<4>[5131059.992356] c00000004accff78 0000000000000000
<4>[5131059.992409] 0000000000000000 0000000000000000
<4>[5131059.992463] 0000000000000000 0000000000000000
<4>[5131059.992514] 0000000000000000 0000000000000000
<4>[5131059.992558] 0000000000000000 ffffffff8021d910
<4>[5131059.992604] 0000000000000000 0000000000000000
<4>[5131059.992653] 00000000ffff22f9 ffffffffc5526c54
<4>[5131059.992700] 0000000300000152 00000000ffff22f9
<4>[5131059.992751] Call Trace:
<4>[5131059.992776] [<ffffffff804b0954>] _spin_lock_irqsave+0xe4/0x200
<4>[5131059.992834] [<ffffffffc7556f00>] ADJ6_ENTRY_DelHashEntry+0x110/0x2e0 [system]
<4>[5131060.038652] [<ffffffffc7604a18>] ND_ENTRY_DeleteNoBindDummy+0x58/0x90 [system]
<4>[5131060.080571] [<ffffffffc7609988>] nd_SingleEvtProc+0x58/0x80 [system]
<4>[5131060.118757] [<ffffffffc75f7e80>] nd_Thread+0x100/0x1e0 [system]
<4>[5131060.151584] [<ffffffff80266ab0>] kthread+0x140/0x150
<4>[5131060.151657] [<ffffffff8021d910>] kernel_thread_helper+0x10/0x20
<4>[5131060.151715]
<4>[5131060.151740]
<4>[5131060.151751] Instruction dump: 1060fe53 00000000 c0820000 <34424000> e0820000 00
0000c0 1000fff5 00000000 c0820000
<0>[5131060.151886] 5:module name: ksplice_lpu_xlp_1649245709_system_old -- module load
address: 0xffffffffe3ef4000
<0>[5131060.151961] 5:module name: ksplice_lpu_xlp_1649245709_system_new -- module load ad
dress: 0xffffffffe3e84000
<0>[5131060.152023] 5:module name: ksplice_lpu_xlp_1649245709 -- module load address: 0xffffffff
e3e60000
<0>[5131060.152085] 5:module name: system -- module load address: 0xffffffffc3068000
<0>[5131060.152140] 5:module name: addon -- module load address: 0xffffffffc0008000

==========display kernel reboot 20 verbose slot 0 ==========
---------- Reboot record 1 ----------
Recorded at : 2023-01-09 10:31:49.773908
Occurred at : 2023-01-09 10:26:13.227958
Reason : 0x2
Thread : kND/1 (TID: 446)
Context : irq context
Slot : 8
Target Slot : 8
Cpu : 0

VCPU ID : 5
Kernel module info : module name (ksplice_lpu_xlp_1649245709_system_old) module address (0xffffffffe3ef4000)
module name (ksplice_lpu_xlp_1649245709_system_new) module address (0xffffffffe3e84000)
module name (ksplice_lpu_xlp_1649245709) module address (0xffffffffe3e60000)
module name (system) module address (0xffffffffc3068000)
module name (addon) module address (0xffffffffc0008000)
Last 5 thread switches : kND/1 (10:25:49.572960)-->
swapper (10:25:49.572995)-->
kND/1 (10:25:49.573096)-->
dipv6_buf (10:25:49.573372)-->
kND/1 (10:25:49.573421)

过程分析

三、问题分析
通过分析代码，判断kND任务处理表项时内部的dummy表项从normal链表移动到nobind链表（有hash锁，无表项锁）时出的问题；
arp线程收到报文之后，会将报文从dummy表改成正常表项（下驱动之后修改，有表项锁，无hash锁；
两个流程并发，但是无锁保护，导致异常；

解决方法

四、H05补丁已经解决