

"KNI 문제 해결: QvPC-DI 플랫폼의 메모리 부족 " 오류

목차

[소개](#)

[배경](#)

[조사 단계](#)

[1단계. 증상 관찰](#)

[KNI: 메모리 부족 로그](#)

[EGTPC 경로 실패](#)

[2단계. DI-Network 상태 저하 확인](#)

[show session recovery status verbose](#)

[show cloud monitor di-network detail](#)

[show cloud monitor controlplane](#)

[show cloud monitor 데이터 플레인](#)

[3단계. Userspace KNI 삭제 확인](#)

[iftask 통계 표시](#)

[4단계. 하드웨어 드라이버 확인](#)

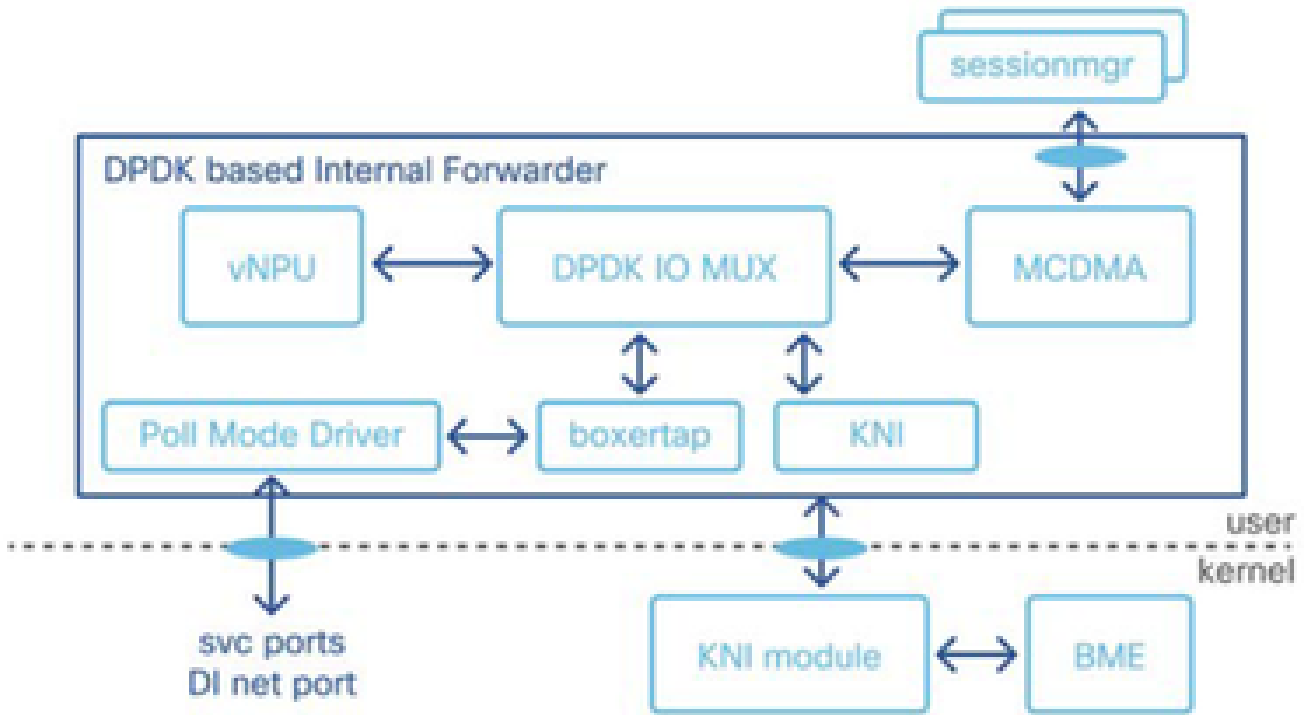
[요약](#)

소개

이 문서에서는 StarOS KNI: 메모리 부족 로그가 StarOS 애플리케이션의 문제 또는 하드웨어 드라이버로 인해 발생하는지 여부를 확인하는 방법에 대해 설명합니다.

배경

DPDK IFTASK(Internal Forwarder) 프로세스 내의 KNI(Kernel Network Interface) 모듈은 사용자 공간 프로그램이 Linux 네트워킹 및 Linux IP 스택을 완전히 우회하여 네트워크 인터페이스에서 직접 패킷을 수신할 수 있도록 하는 메커니즘입니다.



KNI: KNI 모듈에 영향을 주는 리소스 경합 문제가 있을 경우 메모리 부족 로그 속도 제한 경고가 생성됩니다.

1. 메모리 버퍼가 베어 메탈(하드웨어) 레벨에서 지워지지 않으므로 버퍼 오버런이 발생합니다.
2. iftask에서 이러한 패킷에 대한 메시지 버퍼를 할당하는 KNI 풀의 공간이 부족합니다.
3. 가상 함수는 더 많은 패킷을 처리하지만 물리적 함수는 아무 것도 없다고 응답합니다.
4. KNI: 메모리 부족 상태가 발생하면 iftask가 백업 메모리 풀로 이동하여 패킷을 추가로 할당하고 처리합니다. 백업 풀도 메모리가 부족하면 시스템은 패킷을 삭제합니다.
5. iftask에서 커널에서 오는 패킷의 버스트를 읽을 수 없으므로 StarOS에서 KNI: 메모리 부족 로그가 생성됩니다.

KNI에 대한 트리거: 메모리 부족 조건:

버퍼 오버플로 조건에 대한 잠재적인 트리거는 SFTP 또는 SCP 애플리케이션을 실행하거나 CF 및 SF 카드 간에 매우 큰 파일 전송을 실행하는 등 다양할 수 있습니다.

조사 단계

1단계. 증상 관찰

2단계. DI-Network 상태 저하 확인

3단계. Userspace KNI 삭제 확인

4단계. 하드웨어 드라이버 확인

1단계. 증상 관찰

KNI: 메모리 부족 오류를 패킷 손실 또는 애플리케이션 레이어 저하(egtpc 경로 장애)와 같은 다른 증상과 상관시킵니다.

KNI: 메모리 부족 로그

- StarOS Syslogs에서 커널 네트워크 인터페이스의 메모리가 부족함을 나타내는 로그를 볼 수 있습니다.

```
2023-Nov-16+09:18:03.205 [iftask 214701 error] [1/0/9602 <evlogd:0> evlgd_syslogd.c:236] [software inte
```

- 백업 메모리가 모두 소진된 경우 백업 풀의 메모리도 모두 소진되었음을 나타내는 오류 메시지가 표시됩니다.

```
RTE_LOG(ERR, KNI, "Out of memory from Backup pool, kni port %s, socket_id=%d, total=%d, iter=%d\n", kni
```

- 디버그 셸의 tmp 디렉토리에 있는 IFTask 로그에서 KNI: 메모리 부족 오류를 관찰할 수 있습니다.

```
Wed Nov 15 17:20:30 2023 PID:7387 KNI: Out of memory, kni port cpbond0, socket_id=0, total=-759247296,
```

EGTPC 경로 실패

- 여러 피어에 대한 gtpc 경로 실패의 급증은 원인으로 인해 발생할 수 있습니다. 피어에서의 응답은 패킷 손실 중에 발생할 수 없습니다.

```
2023-10-23T00:14:33.813+00:00 Nodename evlogd: [local-60sec33.780] [egtpmgr 143137 info] [6/0/12364 <eg
```

2단계. DI-Network 상태 저하 확인

성능이 저하된 연결을 찾습니다. 지속적으로 볼 때, DI-network 상태 출력의 삭제 또는 손실 비율이 높을수록 DI-network 컨피그레이션 또는 운영 문제, 트래픽 오버로드, VM 또는 호스트 문제를 나타낼 수 있습니다.

show session recovery status verbose

- show session recover status verbose 출력을 사용하여 어떤 가상 기능 카드가 Demux 카드로 사용되고 있는지 확인합니다.

***** show session recovery status verbose *****

Tuesday October 24 11:23:45 EDT 2023

Session Recovery Status:

Overall Status : Ready For Recovery
Last Status Update : 1 second ago

cpu state	----sessmgr---		----aaamgr----		demux active	status
	active	standby	active	standby		
3/0 Active	24	1	24	1	0	Good
4/0 Active	24	1	24	1	0	Good
5/0 Active	24	1	24	1	0	Good
6/0 Active	0	0	0	0	10	Good (Demux)
7/0 Active	24	1	24	1	0	Good
8/0 Active	24	1	24	1	0	Good
9/0 Active	24	1	24	1	0	Good
10/0 Active	24	1	24	1	0	Good
11/0 Active	24	1	24	1	0	Good
12/0 Standby	0	24	0	24	0	Good

show cloud monitor di-network detail

- "show cloud monitor di-network detail" 출력을 사용하여 가상 기능 카드 간 어떤 DI-network 연결이 하트비트에서 삭제되었는지 확인합니다.

- CF 및 SF 카드에서 SF 카드 6으로의 하트비트 드롭이 표시됩니다. 다른 CF 및 SF 카드에 대한 출력에는 하트비트 드롭이 표시되지 않습니다.

***** show cloud monitor di-network detail *****

Tuesday October 24 11:23:51 EDT 2023

Card 1 Heartbeat Results:

ToCard	Health	5Min-Loss	60Min-Loss
6	Good	0.00%	0.66%

Card 2 Heartbeat Results:

6	Bad	14.67%	3.50%
---	-----	--------	-------

Card 3 Heartbeat Results:

6	Bad	5.35%	2.69%
---	-----	-------	-------

Card 4 Heartbeat Results:

6	Good	0.00%	0.00%
---	------	-------	-------

```
...
Card 5 Heartbeat Results:
...
  6      Bad      18.57%      3.90%
```

```
...
Card 6 Heartbeat Results:
...
  1      Good      0.00%      0.90%
  2      Bad      12.63%     3.31%
  3      Bad      2.90%      2.14%
  4      Good      0.00%      0.00%
  5      Bad      13.09%     3.30%
  7      Good      0.00%      0.00%
  8      Bad      2.91%      2.20%
  9      Good      0.00%      0.93%
 10      Bad      14.28%     3.38%
 11      Bad      3.67%      2.09%
 12      Good      0.00%      0.00%
```

```
...
Card 7 Heartbeat Results:
...
  6      Good      0.00%      0.00%
```

```
...
Card 8 Heartbeat Results:
...
  6      Bad      7.47%      2.85%
```

```
...
Card 9 Heartbeat Results:
...
  6      Bad      0.00%      1.07%
```

```
...
Card 10 Heartbeat Results:
...
  6      Bad      16.01%     3.73%
```

```
...
Card 11 Heartbeat Results:
...
  6      Bad      7.47%      2.71%
```

```
...
Card 12 Heartbeat Results:
...
  6      Good      0.00%      0.00%
```

show cloud monitor controlplane

- show cloud monitor 컨트롤플레인 출력을 사용하여 성능 저하가 있는 DI 네트워크 연결을 식별합니다.

```
***** show cloud monitor controlplane *****
Tuesday October 24 11:24:22 EDT 2023
```

Cards		15 Second Interval			5 Minute Interval			60 Minute Interval		
Src	Dst	Xmit	Recv	Miss%	Xmit	Recv	Miss%	Xmit	Recv	Miss%
01	06	75	75	0.0%	1500	1500	0.0%	18000	17842	0.9%

...	02	06	75	75	0.0%	1500	1265	15.7%	18000	17546	2.5%
...	03	06	75	75	0.0%	1500	1396	6.9%	18000	17491	2.8%
...	04	06	75	75	0.0%	1500	1500	0.0%	18000	18000	0.0%
...	05	06	75	75	0.0%	1500	1267	15.5%	18000	17325	3.8%
...	06	01	75	75	0.0%	1500	1500	0.0%	18000	17823	1.0%
...	06	02	75	75	0.0%	1500	1301	13.3%	18000	17567	2.4%
...	06	03	75	75	0.0%	1500	1419	5.4%	18000	17561	2.4%
...	06	04	75	75	0.0%	1500	1500	0.0%	18000	18000	0.0%
...	06	05	75	75	0.0%	1500	1294	13.7%	18000	17579	2.3%
...	06	07	75	75	0.0%	1500	1500	0.0%	18000	18000	0.0%
...	06	08	75	75	0.0%	1500	1417	5.5%	18000	17565	2.4%
...	06	09	75	75	0.0%	1500	1500	0.0%	18000	17824	1.0%
...	06	10	75	75	0.0%	1500	1296	13.6%	18000	17573	2.4%
...	06	11	75	75	0.0%	1500	1422	5.2%	18000	17570	2.4%
...	06	12	75	75	0.0%	1500	1500	0.0%	18000	18000	0.0%
...	07	06	75	75	0.0%	1500	1500	0.0%	18000	18000	0.0%
...	08	06	75	75	0.0%	1500	1426	4.9%	18000	17545	2.5%
...	09	06	75	75	0.0%	1500	1500	0.0%	18000	17833	0.9%
...	10	06	75	75	0.0%	1500	1278	14.8%	18000	17369	3.5%
...	11	06	75	75	0.0%	1500	1408	6.1%	18000	17481	2.9%
...	12	06	75	75	0.0%	1500	1500	0.0%	18000	18000	0.0%

show cloud monitor 데이터 플레인

- show cloud monitor dataplane 출력을 사용하여 어떤 DI-네트워크 연결에 디그레이션이 있는지 확인하고 가상 기능 카드 간의 단방향 디그레이션을 확인합니다.

***** show cloud monitor dataplane *****

Tuesday October 24 11:21:46 EDT 2023

Cards		15 Second Interval			5 Minute Interval			60 Minute Interval			
Src	Dst	Miss	Hit	Pct	Miss	Hit	Pct	Miss	Hit	Pct	
...	06	01	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...	06	02	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...	06	03	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...	06	04	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...	06	05	1	149	0.7%	0	3001	0.0%	0	36000	0.0%
...	01	06	0	150	0.0%	0	3000	0.0%	0	36000	0.0%

02	06	0	150	0.0%	210	2790	7.0%	1015	34985	2.8%
03	06	31	119	20.7%	540	2460	18.0%	995	35005	2.8%
04	06	34	116	22.7%	554	2446	18.5%	1017	34983	2.8%
05	06	0	150	0.0%	213	2787	7.1%	991	35009	2.8%
07	06	0	150	0.0%	0	3000	0.0%	359	35641	1.0%
08	06	29	121	19.3%	546	2454	18.2%	1009	34991	2.8%
09	06	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
10	06	0	150	0.0%	208	2792	6.9%	992	35008	2.8%
11	06	31	119	20.7%	548	2452	18.3%	993	35007	2.8%
12	06	34	116	22.7%	547	2453	18.2%	1001	34999	2.8%
...										
06	07	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...										
06	08	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...										
06	09	0	150	0.0%	0	3000	0.0%	1	35999	0.0%
...										
06	10	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...										
06	11	0	150	0.0%	0	3000	0.0%	0	36000	0.0%
...										
06	12	0	150	0.0%	0	3000	0.0%	0	36000	0.0%

3단계. Userspace KNI 삭제 확인

iftask 통계 표시

- show iftask stats 출력을 여러 번 수집하여 KNI 삭제로 인해 IFTASK userspace 애플리케이션 레벨(StarOS)이 증가하지 않는지 확인합니다.

```
***** show iftask stats *****
Tuesday October 24 11:22:06 EDT 2023
```

```
...
```

CARD 6 STATS			
Counters	SF6	SF6_PPS	
svc_rx	2587301598	2203	
svc_tx	548969428	295	
di_rx	2260147059	2258	
di_tx	4072038717	3966	
__ALL_DROPS__	0	0	
svc_tx_drops	0	0	
di_rx_drops	0	0	
di_tx_drops	0	0	
sw_rss_enq_drops	0	0	
kni_thread_drops	0	0	
kni_drops	0	0	
mcdma_drops	0	0	
mux_deliver_hop_drops	0	0	
mux_deliver_drops	0	0	
mux_xmit_failure_drops	0	0	
mc_dma_thread_enq_drops	0	0	
sw_tx_egress_enq_drops	0	0	
cpeth0_drops	0	0	
mcdma_summary_drops	0	0	

```

fragmentation_err          0          0
reassembly_err             0          0
reassembly_ring_enq_err    0          0
__DISCARDS__               241984     0
__BOND_DISCARDS__         55282718  142
...

```

TOTAL STATS

```

-----
Counters          TOTAL          TOTAL_PPS
-----
svc_rx            27964563261      24791
svc_tx            36109966153      30168
di_rx             74133486629      51929
di_tx             73958155063      50897
__ALL_DROPS__    0                0
svc_tx_drops     0                0
di_rx_drops     0                0
di_tx_drops     0                0
sw_rss_enq_drops 0                0
kni_thread_drops 0                0
kni_drops       0                0
mcdma_drops     0                0
mux_deliver_hop_drops 0            0
mux_deliver_drops 0            0
mux_xmit_failure_drops 0            0
mc_dma_thread_enq_drops 0            0
sw_tx_egress_enq_drops 0            0
cpeth0_drops    0                0
mcdma_summary_drops 0            0
fragmentation_err 0                0
reassembly_err   0                0
reassembly_ring_enq_err 0            0
__DISCARDS__    2324968         0
__BOND_DISCARDS__ 55635534       149
-----

```

```

NDR is          100.0000
CONTINUE_TRAFFIC
-----

```

4단계. 하드웨어 드라이버 확인

애플리케이션 레이어에서 결함이 제거된 상태에서 하드웨어 레벨의 기본 드라이버에 초점을 맞추어 KNI: 메모리 부족 오류를 해결합니다.

베어 메탈 하드웨어 드라이버가 각 가상 기능에 대해 일정량의 버퍼를 할당하기 때문에 리소스 결합 문제는 일반적으로 하드웨어 레벨에서 드라이버가 일치하지 않거나 결함이 있는 드라이버로 인해 발생합니다. 응용 프로그램에 필요한 버퍼를 할당한 결함이 있는 하드웨어 드라이버가 메모리를 해제하지 않았습니다.

타사(타사) 가상화 소프트웨어 및/또는 하드웨어를 사용 중인 경우, 버전 및 드라이버에서 호환성 불일치 또는 결함이 있는지 조사하십시오.

요약

KNI: 메모리 부족 오류가 애플리케이션 수준 프로세스 또는 기본 하드웨어 드라이버에 의해 발생하는지 확인하려면 DI-네트워크 성능 저하 및 사용자 공간 KNI 삭제의 증거를 확인하십시오. 해당 사용자공간 KNI 저하 없이 DI-network 열화가 존재하면 그 원인은 하드웨어 수준으로 결론 내릴 수 있다. KNI: 하드웨어 수준 삭감과 함께 메모리 부족 오류가 발생하면 하드웨어 드라이버에 오류가 발생한 것입니다.

영향을 받는 애플리케이션 레벨 StarOS 가상 함수가 상주하는 노드를 오프로드하고 호스트를 다시 로드하면 기본 컴퓨팅의 메모리 버퍼가 일시적으로 지워져 오류 및 패킷 손실이 일시적으로 감소할 수 있습니다. 그러나 이는 영구적인 해결책이 아닙니다! 패킷 손실 및 KNI: 결함이 있는 하드웨어 드라이버에서 버퍼 오버플로 조건이 재발할 때 메모리 부족 오류가 다시 발생합니다.

이 번역에 관하여

Cisco는 전 세계 사용자에게 다양한 언어로 지원 콘텐츠를 제공하기 위해 기계 번역 기술과 수작업 번역을 병행하여 이 문서를 번역했습니다. 아무리 품질이 높은 기계 번역이라도 전문 번역가의 번역 결과물만큼 정확하지는 않습니다. Cisco Systems, Inc.는 이 같은 번역에 대해 어떠한 책임도 지지 않으며 항상 원본 영문 문서(링크 제공됨)를 참조할 것을 권장합니다.