

"KNI:QvPC-DIプラットフォームでのメモリ不足"; エラーのトラブルシューティング

内容

[はじめに](#)

[背景](#)

[調査手順](#)

[ステップ 1: 症状の確認](#)

[KNI: メモリ不足ログ](#)

[EGTPCバスの障害](#)

[ステップ 2: DIネットワークの健全性低下のチェック](#)

[show session recovery status verbose \(詳細を表示\)](#)

[show cloud monitor di-network detail \(クラウドモニタdiネットワークの詳細を表示\)](#)

[クラウドモニタコントロールプレーンの表示](#)

[クラウドモニタデータプレーンの表示](#)

[ステップ 3: ユーザ空間のKNIドロップの確認](#)

[show iftask statsの実行](#)

[ステップ 4: ハードウェアドライバの確認](#)

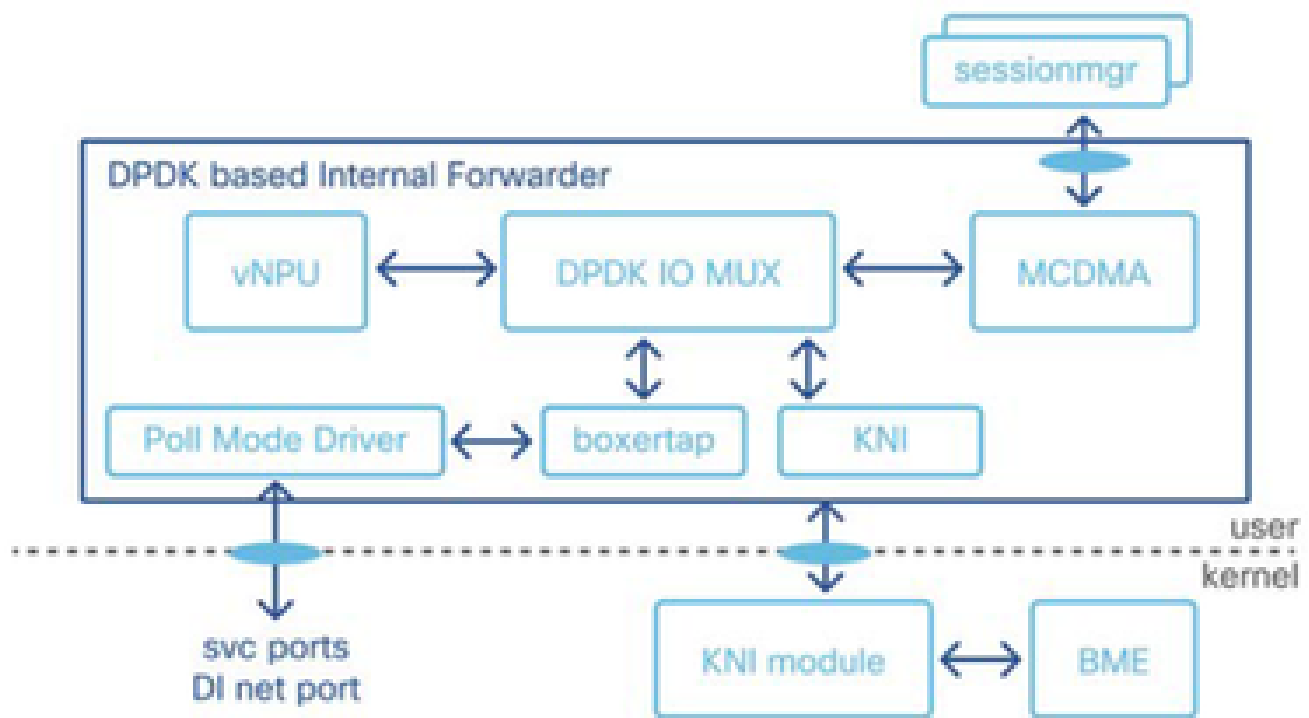
[要約](#)

はじめに

このドキュメントでは、StarOS KNI:Out of MemoryログがStarOSアプリケーションの問題またはハードウェアドライバによって発生しているかどうかを判別する方法について説明します。

背景

DPDK内部フォワーダ(IFTASK)プロセス内のカーネルネットワークインターフェイス(KNI)モジュールは、ユーザ空間のプログラムがネットワークインターフェイスから直接パケットを受信できるようにし、LinuxネットワークとLinux IPスタックを完全にバイパスするメカニズムです。



KNI：メモリ不足ログKNIモジュールに影響するリソース競合の問題が発生すると、レート制限の警告が表示されます。

1. メモリバッファはベアメタル（ハードウェア）レベルではクリアされず、バッファのオーバーランを引き起こします。
2. iftaskがこれらのパケット用にメッセージバッファを割り当てるKNIプールが容量不足になります。
3. 仮想関数は他のパケットを照会しますが、物理関数は何も持っていないことを応答します。
4. KNI:Out of Memory状態が発生すると、iftaskはバックアップメモリプールに入り、パケットをさらに割り当てて処理します。バックアッププールのメモリも不足すると、システムはパケットをドロップします。
5. iftaskではカーネルからのパケットのバーストを読み取ることができないので、StarOSでKNI: Out of Memoryログが生成されます。

KNI：メモリ不足状態のトリガー:

バッファオーバーフロー状態の潜在的なトリガーは、SFTPまたはSCPアプリケーションの実行、CFカードとSFカード間の非常に大きなファイル転送など、さまざまです。

調査手順

ステップ 1：症状の確認

ステップ 2：DIネットワークの健全性低下のチェック

ステップ 3：ユーザ空間のKNIドロップの確認

ステップ 4 : ハードウェアドライバの確認

ステップ 1 : 症状の確認

KNI:Out of Memoryエラーのタイミングを、パケット損失やアプリケーション層の劣化(egtpcパス障害)などの他の症状と関連付けます。

KNI : メモリ不足ログ

- StarOSのSyslogでは、カーネルのネットワークインターフェイスがメモリ不足であることを示すログが表示されます。

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

- バックアップメモリを使い果たした場合、バックアッププールのメモリも使い果たされたことを示すエラーメッセージが表示されます。

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

- デバッグシエルのtmpディレクトリにあるIFTaskログで、「KNI: Out of Memory」エラーを確認できます。

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

EGTPCパスの障害

- さまざまなピアへのgtpcパス障害でスパイクが発生する可能性があります。その原因は、パケット損失時にピアからの応答がないことにあります。

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

ステップ 2 : DIネットワークの健全性低下のチェック

パフォーマンスが低下している接続を特定します。継続的に見ると、DIネットワークの健全性出力の低下または損失率が高い場合は、DIネットワークの設定または動作の問題、トラフィックの過負荷、またはVMまたはホストの問題を示している可能性があります。

show session recovery status verbose (詳細を表示)

- show session recover status verbose出力を使用して、Demuxカードとして機能している仮想機能カードを識別します。

```
***** 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 (クラウドモニタDIネットワークの詳細を表示)

- 「show cloud monitor di-network detail」の出力を使用して、バーチャルファンクションカード間のどのDIネットワーク接続でハートビートのドロップが発生しているかを特定します。

- CFカードおよびSFカードからSFカード6へのハートビートでのドロップが示されています。他のCFカードおよびSFカードに対するCFカードおよびSFカードの出力では、ハートビートドロップが表示されません。

```
***** 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%

```

クラウドモニタコントロールプレーンの表示

- show cloud monitor controlplaneの出力を使用して、どのDIネットワーク接続のパフォーマンスが低下しているかを識別します。

***** 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%

クラウドモニタデータプレーンの表示

- show cloud monitor dataplaneの出力を使用して、どのDIネットワーク接続に品質低下があるかを識別し、仮想機能カード間の単方向品質低下を特定します。

***** 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%

ステップ 3 : ユーザ空間のKNIドロップの確認

show iftask statsの実行

- show iftask stats出力を複数回収集して、IFTASKのユーザ空間のアプリケーションレベル (StarOS)でKNIドロップが増加していないことを確認します。

```
***** 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
-----

```

ステップ 4 : ハードウェアドライバの確認

アプリケーション層が原因ではないことが明らかになったら、KNI:Out of Memoryエラーに対処するために、ハードウェアレベルの基盤となるドライバに焦点を当てます。

ベアメタルハードウェアドライバは仮想関数ごとに一定量のバッファを割り当てるため、リソース競合の問題は通常、ドライバの不一致やハードウェアレベルでのドライバの欠陥が原因で発生します。アプリケーションに必要なバッファを割り当てた不良ハードウェアドライバは、メモリを解放しませんでした。

サードパーティ（シスコ以外）の仮想化ソフトウェアやハードウェアを使用している場合は、バージョンやドライバに互換性の不一致や不具合がないかどうかを調べます。

要約

KNI:Out of Memoryエラーの原因が、アプリケーションレベルのプロセスまたは基盤となるハードウェアドライバにあるかどうかを判断するには、DIネットワークの品質低下とユーザ領域のKNIドロップの証拠をチェックします。対応するユーザ空間KNIデグラデーションなしでDIネットワークの劣化が存在する場合、原因はハードウェアレベルにあると結論付けることができます。KNI:ハードウェアレベルの低下を伴うメモリ不足エラーは、ハードウェアドライバの障害を示します。

影響を受けるアプリケーションレベルのStarOS仮想関数が存在するノードのオフロードとホストの計算のリロードによって、基盤となる計算上のメモリバッファが一時的にクリアされ、エラーとパケット損失が一時的に減少する可能性があります。ただし、これは恒久的な解決策ではありません。パケット損失とKNI:Out of Memoryエラーが再発します。これは、障害のあるハードウェアドライバでバッファオーバーフロー状態が再発した場合です。

翻訳について

シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。