

Quantum 虛擬化封包核心分散式例項(QvPC-DI):iftask和網路處理器單元(NPU)效能監控

目錄

[簡介](#)

[採用元件](#)

[Iftask體系結構](#)

[監視檔案效能](#)

[批次統計](#)

簡介

本文檔介紹如何在QvPC-DI上監控iftask / NPU的效能。

採用元件

本文檔中的資訊基於QvPC-DI。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除 (預設) 的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

Iftask體系結構

iftask是QvPC-DI中的一個進程。它為DI網路埠和服務埠在服務功能虛擬卡(SF)和控制功能虛擬卡(CF)上啟用資料平面開發套件(DPDK)功能。DPDK是處理虛擬化環境中輸入/輸出的更有效方式。

高效能網路介面控制器(NIC)的裝置驅動程式現在被移至userspace，以避免昂貴的情景交換機(userspace/kernelspace)。

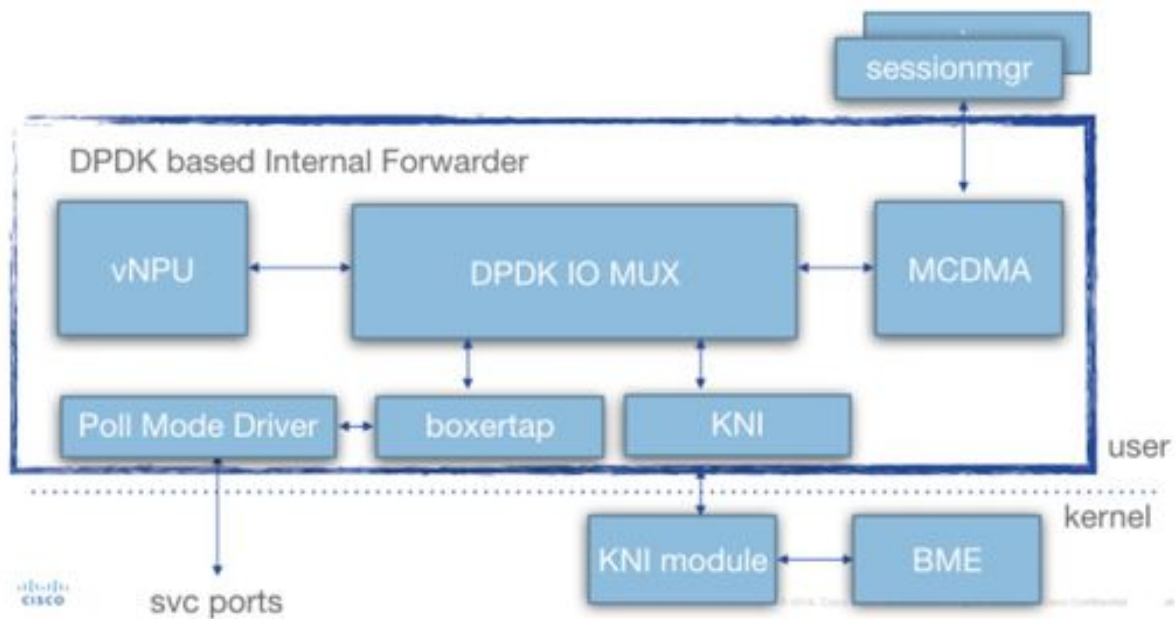
驅動程式在使用者空間中以不可中斷模式運行，執行緒可以直接訪問這些NIC驅動程式中的硬體隊列/環緩衝區。

有關架構的文檔可從以下網址獲得：

[Ultra Gateway Platform System Administration Guide](#)中的Ultra Services Platform(USP)簡介。

[不同版本的可用性。](#)

深入的iftask體系結構 (用於SF) 如下圖所示：



在部署期間，將一定數量的虛擬中央處理器(vCPU)靜態分配給iftask進程。這減少了用於使用者空間應用程式 (sessmgr等) 的核心數量，但大大提高了I/O的效能。

通過在部署期間與每個SF/CF關聯的param.cfg模板中的以下引數完成此分配：

- IFTASK_CORES (將分配給iftask的可用核心的百分比)
- (IFTASK_CRYPTO_CORES — 此處未討論)
- (IFTASK_MCDMA_CORES — 此處未討論)

1. 在SF上，iftask進程在內部將其分配的核心分配到：輪詢模式驅動程式(PMD)vCPU (執行 tx/rx/vnpu活動) MCDMA vCPU，執行從iftask到sessmgr和回傳的資料包傳輸
2. 在CF上，不需要MCDMA vCPU，因為SF不承載sessmgr進程。

show cloud hardware iftask命令提供有關您的QVPC-DI部署的詳細資訊：

```
[local]UGP# show cloud hardware iftask
Card 1:
  Total number of cores on VM:      8
  Number of cores for PMD only:     0
  Number of cores for VNPU only:    0
  Number of cores for PMD and VNPU: 2    <-- CF: 2 out of 8 cores are assigned to iftask
PMD/VNPU
  Number of cores for MCDMA:        0    <-- CF: no cores allocated to MCDMA as there is no
sessmgr process on CF
  Number of cores for Crypto:       0
  Hugepage size:                    2048 kB
  Total hugepages:                  3670016 kB
  NPUSHM hugepages:                 0 kB
  CPU flags: avx sse sse2 ssse3 sse4_1 sse4_2
  Poll CPU's: 1 2
  KNI reschedule interval: 5 us
...
Card 3:
  Total number of cores on VM:      8
  Number of cores for PMD only:     0
  Number of cores for VNPU only:    0
```

```
Number of cores for PMD and VNPU: 2 <-- SF: 2 out of 8 core are assigned to iftask
PMD/VNPU
Number of cores for MCDMA: 1 <-- SF: 1 out of 8 cores is assigned to iftak MCDMA
Number of cores for Crypto: 0
Hugepage size: 2048 kB
Total hugepages: 4718592 kB
NPUSHM hugepages: 0 kB
CPU flags: avx sse sse2 ssse3 sse4_1 sse4_2
Poll CPU's: 1 2 3
KNI reschedule interval: 5 us
```

命令「**show cloud configuration**」將提供有關所用引數的更多詳細資訊：

```
[local]UGP# show cloud configuration
Card 1:
  Config Disk Params:
  -----
CARDSLOT=1
CPUID=0
CARDTYPE=0x40010100
DI_INTERFACE=BOND:TYPE:ixgbevf-1,TYPE:ixgbevf-2
DI_INTERFACE_VLANID=2111
VNFM_INTERFACE=MAC:fa:16:3e:23:aa:e9
VNFM_PROXY_ADDRS=172.16.180.3,172.16.180.5,172.16.180.6
MGMT_INTERFACE=MAC:fa:16:3e:87:23:9b
VNFM_IPV4_ENABLE=true
VNFM_IPV4_DHCP_ENABLE=true

  Local Params:
  -----
CARDSLOT=1
CARDTYPE=0x40010100
CPUID=0
...
Card 3:
  Config Disk Params:
  -----
CARDSLOT=3
CPUID=0
CARDTYPE=0x42030100
DI_INTERFACE=BOND:TYPE:ixgbevf-1,TYPE:ixgbevf-2
SERVICE1_INTERFACE=BOND:TYPE:ixgbevf-3,TYPE:ixgbevf-4
SERVICE2_INTERFACE=BOND:TYPE:ixgbevf-5,TYPE:ixgbevf-6
DI_INTERFACE_VLANID=2111
VNFM_INTERFACE=MAC:fa:16:3e:29:c6:b7
IFTASK_CORES=30
VNFM_IPV4_ENABLE=true
VNFM_IPV4_DHCP_ENABLE=true

  Local Params:
  -----
CARDSLOT=3
CARDTYPE=0x42010100
CPUID=0
```

監視檔案效能

iftask過程可以通過多種方式監控。

整合show命令清單：

```

show subscribers data-rate
show npumgr dinet utilization pps
show npumgr dinet utilization pps
show cloud monitor di-network summary
show cloud hardware iftask
show cloud configuration
show iftask stats summary
show port utilization table
show npu utilization table
show npumgr utilization information
show processes cpu

```

命令 `#show cpu info verbose` 不會提供有關 iftask 核心的資訊。它們將始終以 100% 利用率列出。

在下面的示例中，核心 1、2、3 與 iftask 關聯，並且列出的利用率是 100%，這是預期的。

```

Card 3, CPU 0:
  Status                : Standby, Kernel Running, Tasks Running
  Load Average          : 3.12, 3.12, 3.13 (3.95 max)
  Total Memory           : 16384M
  Kernel Uptime         : 4D 21H 56M
  Last Reading:
    CPU Usage All       : 1.9% user, 0.3% sys, 0.0% io, 0.0% irq, 97.8% idle
    Core 0              : 5.8% user, 0.2% sys, 0.0% io, 0.0% irq, 94.0% idle
    Core 1              : Not Averaged (Poll CPU)
    Core 2              : Not Averaged (Poll CPU)
    Core 3              : Not Averaged (Poll CPU)
    Core 4              : 2.2% user, 0.2% sys, 0.0% io, 0.0% irq, 97.6% idle
    Core 5              : 0.8% user, 0.5% sys, 0.0% io, 0.0% irq, 98.7% idle
    Core 6              : 0.4% user, 0.5% sys, 0.0% io, 0.0% irq, 99.1% idle
    Core 7              : 0.1% user, 0.3% sys, 0.0% io, 0.0% irq, 99.6% idle
  Poll CPUs             : 3 (1, 2, 3)
    Core 1              : 100.0% user, 0.0% sys, 0.0% io, 0.0% irq, 0.0% idle
    Core 2              : 100.0% user, 0.0% sys, 0.0% io, 0.0% irq, 0.0% idle
    Core 3              : 100.0% user, 0.0% sys, 0.0% io, 0.0% irq, 0.0% idle
  Processes / Tasks     : 143 processes / 16 tasks
  Network mcdmaN        : 0.002 kpps rx, 0.001 mbps rx, 0.002 kpps tx, 0.001 mbps tx
  File Usage            : 1504 open files, 1627405 available
  Memory Usage          : 7687M 46.9% used
  Memory Details:
    Static              : 330M kernel, 144M image
    System              : 10M tmp, 0M buffers, 54M kcache, 79M cache
    Process/Task        : 6963M (120M small, 684M huge, 6158M other)
    Other               : 104M shared data
    Free                : 8696M free
    Usable              : 5810M usable (8696M free, 0M reclaimable, 2885M reserved by tasks)

```

命令 `#show npu` 使用率表將很好地總結與 iftask 過程（在每個卡上）關聯的每個核心的利用率。

附註：這裡的重要一點是，確定某些核心的利用率是否始終高於其他核心。

```

[local]UGP# show npu utilization table
-----iftask-----
lcore   now    5min  15min
-----
01/0/1   0%     0%    0%
01/0/2   0%     0%    0%

```

```

02/0/1      0%    0%    0%
02/0/2      2%    1%    0%
03/0/1      0%    0%    0%
03/0/2      0%    0%    0%
03/0/3      0%    0%    0%
04/0/1      0%    0%    0%
04/0/2      0%    0%    0%
04/0/3      0%    0%    0%
05/0/1      0%    0%    0%
05/0/2      0%    0%    0%
05/0/3      0%    0%    0%

```

命令#show npumgr utilization information(隱藏命令)

此命令可提供有關每個完整核心的詳細資訊，以及這些核心上佔用CPU的資源的詳細資訊。

附註：PMD核心的CPU消耗在PortRX、PortTX、KNI、密碼上。

附註：MCDMA核心的CPU被MCDMA佔用。

PMD和MCDMA核心負載均勻。

如果不是這樣，則可能需要進行一些調整（例如，分配更多/更少的MDMA核心）。

```

***** show npumgr utilization information 3/0/0 *****
 5-Sec Avg: lcore01| lcore02| lcore03| lcore04| lcore05| lcore06| lcore07| lcore08| lcore09|
lcore10| lcore11| lcore12|
      Idle:      41%|    47%|    73%|    62%|    48%|    49%|    69%|    71%|    49%|
64%|    48%|    69%|
      PortRX:    32%|    36%|    0%|    0%|    0%|    0%|    31%|    29%|    0%|
0%|    0%|    31%|
      PortTX:    16%|    18%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|    0%|    0%|
      KniRX:     3%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|    0%|    0%|
      McdmaRX:   0%|    0%|    11%|    15%|    22%|    21%|    0%|    0%|    20%|
14%|    20%|    0%|
      Mcdma:     0%|    0%|    1%|    2%|    3%|    2%|    0%|    0%|    3%|
3%|    3%|    0%|
      McdmaFlush: 0%|    0%|    15%|    21%|    27%|    27%|    0%|    0%|    27%|
18%|    28%|    0%|
      Cipher:    7%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|    0%|    0%|
      rx kbits/sec: 173914| 257249| 248543| 224081| 319973| 299437| 450992| 380112| 391400|
325099| 355809| 399515|
      rx frames/sec: 30557| 36549| 37465| 32560| 46914| 43938| 70711| 54818| 58482|
47462| 53477| 58793|
      tx kbits/sec: 169641| 251958| 253930| 228760| 326714| 305750| 440773| 372187| 399806|
331914| 363491| 391002|
      tx frames/sec: 30551| 36524| 37465| 32560| 46914| 43938| 70681| 54785| 58483|
47462| 53477| 58749|

 5-Min Avg: lcore01| lcore02| lcore03| lcore04| lcore05| lcore06| lcore07| lcore08| lcore09|
lcore10| lcore11| lcore12|
      Idle:      18%|    36%|    60%|    62%|    46%|    45%|    65%|    62%|    44%|
53%|    39%|    65%|
      PortRX:    29%|    45%|    0%|    0%|    0%|    0%|    35%|    38%|    0%|
0%|    0%|    35%|

```

```

PortTX:      17%|    20%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
KnirRX:      3%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
Kni:         1%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
McdmaRX:     0%|    0%|   17%|   17%|   23%|   24%|    0%|    0%|    0%|   24%|
19%|        26%|    0%|
Mcdma:       0%|    0%|    2%|    2%|    3%|    3%|    0%|    0%|    0%|    4%|
3%|         4%|    0%|
McdmaFlush:  0%|    0%|   21%|   20%|   28%|   28%|    0%|    0%|    0%|   28%|
24%|        30%|    0%|
Cipher:      32%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
rx kbits/sec: 217296| 297699| 261605| 268546| 389380| 397101| 479195| 528945| 438931|
360583| 454262| 468575|
rx frames/sec: 39194| 42022| 40122| 40086| 58217| 59507| 80931| 76094| 67224|
54592| 68565| 67013|
tx kbits/sec: 211773| 291616| 267373| 274308| 397747| 405655| 467493| 517944| 448590|
368412| 464116| 458868|
tx frames/sec: 39182| 41998| 40122| 40086| 58217| 59507| 80895| 76058| 67224|
54592| 68565| 66973|

```

```

15-Min Avg: lcore01| lcore02| lcore03| lcore04| lcore05| lcore06| lcore07| lcore08| lcore09|
lcore10| lcore11| lcore12|
Idle:        22%|   37%|   60%|   61%|   45%|   46%|   64%|   64%|   44%|
49%|         41%|   66%|
PortRX:      33%|   43%|    0%|    0%|    0%|    0%|   36%|   36%|    0%|
0%|          0%|  34%|
PortTX:      18%|   20%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
KnirRX:      4%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
McdmaRX:     0%|    0%|   16%|   16%|   24%|   24%|    0%|    0%|    0%|   24%|
22%|        25%|    0%|
Mcdma:       0%|    0%|    2%|    2%|    3%|    3%|    0%|    0%|    0%|    4%|
4%|         4%|    0%|
McdmaFlush:  0%|    0%|   21%|   20%|   28%|   28%|    0%|    0%|    0%|   28%|
26%|        30%|    0%|
Cipher:      23%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|    0%|
0%|          0%|    0%|
rx kbits/sec: 225682| 297062| 267631| 272770| 419821| 405286| 489633| 511100| 476834|
401810| 467134| 466549|
rx frames/sec: 39670| 42772| 40892| 40834| 62740| 61170| 83540| 76519| 72158|
60242| 70707| 67961|
tx kbits/sec: 220089| 290875| 273510| 278639| 428840| 414080| 477557| 500037| 487201|
410421| 477298| 456711|
tx frames/sec: 39657| 42748| 40892| 40834| 62740| 61170| 83504| 76484| 72158|
60242| 70707| 67925|

```

@ tick 896633 (+ve-skew-cnt=123633, -ve-skew-cnt=0), failed samples 0

命令#show npumgr dinet utilization pps、#show dinet utilization bps和#show port utilization table

它們提供有關DI埠和服務埠上的負載的資訊。

實際效能取決於NIC/CPU和CPU分配給iftask。

```
[local]UGP# show npumgr dinet utilization pps
```

```

----- Average DINet Port Utilization (in kpps) -----
Port  Type                Current          5min          15min
      Rx      Tx      Rx      Tx      Rx      Tx

```

```
-----
```

1/0	Virtual Ethernet	0	0	0	0	0	0
2/0	Virtual Ethernet	0	0	0	0	0	0
3/0	Virtual Ethernet	0	0	0	0	0	0
4/0	Virtual Ethernet	0	0	0	0	0	0
5/0	Virtual Ethernet	0	0	0	0	0	0

[local]UGP# **show npumgr dinet utilization bps**

```
----- Average DInet Port Utilization (in mbps) -----
```

Port	Type	Current		5min		15min	
		Rx	Tx	Rx	Tx	Rx	Tx
1/0	Virtual Ethernet	1	1	1	1	1	1
2/0	Virtual Ethernet	1	0	1	0	1	0
3/0	Virtual Ethernet	0	0	0	0	0	0
4/0	Virtual Ethernet	0	0	0	0	0	0
5/0	Virtual Ethernet	0	0	0	0	0	0

[local]UGP# **show port utilization table**

```
----- Average Port Utilization (in mbps) -----
```

Port	Type	Current		5min		15min	
		Rx	Tx	Rx	Tx	Rx	Tx
1/1	Virtual Ethernet	0	0	0	0	0	0
2/1	Virtual Ethernet	0	0	0	0	0	0
3/10	Virtual Ethernet	0	0	0	0	0	0
3/11	Virtual Ethernet	0	0	0	0	0	0
4/10	Virtual Ethernet	0	0	0	0	0	0
4/11	Virtual Ethernet	0	0	0	0	0	0
5/10	Virtual Ethernet	0	0	0	0	0	0
5/11	Virtual Ethernet	0	0	0	0	0	0

命令#show cloud monitor di-network summary

此命令可監控DI網路的運行狀況。賀卡正在互相傳送心跳，和損失情況。在正常的系統中，不會報告任何損失。

[local]UGP# **show cloud monitor di-network summary**

Card 3 Heartbeat Results:

ToCard	Health	5MinLoss	60MinLoss
1	Good	0.00%	0.00%
2	Good	0.00%	0.00%
4	Good	0.00%	0.00%
5	Good	0.00%	0.00%

Card 4 Heartbeat Results:

ToCard	Health	5MinLoss	60MinLoss
1	Good	0.00%	0.00%
2	Good	0.00%	0.00%
3	Good	0.00%	0.00%
5	Good	0.00%	0.00%

Card 5 Heartbeat Results:

ToCard	Health	5MinLoss	60MinLoss
1	Good	0.00%	0.00%
2	Good	0.00%	0.00%
3	Good	0.00%	0.00%
4	Good	0.00%	0.00%

命令#show iftask stats summary

在NPU負載較高的情況下，流量可能會被丟棄。

要對此進行評估，可以採取命令#show iftask stats summary輸出。

附註：DISCARDS可以是非零值。

附註：所有其他計數器不應遞增。

[local]VPC# show iftask stats summary^M
Thursday January 18 16:01:29 IST 2018

Counter	SF3	SF4	SF5	SF6	SF7
SF8	SF9	SF10	SF11	SF12	___TOTAL___
svc_rx	32491861127	16545600654	37041906441	37466889835	32762859630
34931554543	38861410897	16025531220	33566817747	32823851780	312518283874
svc_tx	46024774071	14811663244	40316226774	39926898585	40803541378
48718868048	35252698559	1738016438	4249156512	40356388348	312198231957
di_rx	42307187425	14637310721	40072487209	39584697117	41150445596
44534022642	31867253533	1731310419	4401095653	40711142205	300996952520
di_tx	28420090751	16267050562	36423298668	36758561246	32731606974
30366650898	35201117980	16009902791	33536789041	32815316570	298530385481
__ALL_DROPS__	1932492	252	17742	790473	11228
627018	844812	60402	0	460830	4745249
svc_tx_drops	0	0	0	0	0
0	0	0	0	0	0
di_rx_drops	0	0	1	0	49
113	579	30200	0	4888	35830
di_tx_drops	0	0	0	0	0
0	0	0	0	0	0
sw_rss_enq_drops	0	0	0	0	0
0	0	0	0	0	0
kni_thread_drops	0	0	0	0	0
0	0	0	0	0	0
kni_drops	0	1	0	0	0
0	124	30200	0	0	30325
mcdma_drops	0	0	0	168	80
194535	758500	0	0	11628	964911
mux_deliver_hop_drops	0	0	0	0	0
0	0	0	1019	1019	0
mux_deliver_drops	0	0	0	0	0
0	0	0	0	0	0
mux_xmit_failure_drops	0	0	3	0	0
0	7	2	0	0	12
mc_dma_thread_enq_drops	0	0	0	0	49
113	580	0	0	3457	4199
sw_tx_egress_enq_drops	1904329	0	0	0	787971
429214	85022	0	0	429810	3645350
cpeth0_drops	0	0	0	0	0
0	0	0	0	0	0
mcdma_summary_drops	28163	247	17742	2334	2046
3043	0	0	10028	63603	0
fragmentation_err	0	0	0	0	0
0	0	0	0	0	0
reassembly_err	0	0	0	0	0
0	0	0	0	0	0
reassembly_ring_enq_err	0	0	0	0	0
0	0	0	0	0	0
__DISCARDS__	20331090	9051092	23736055	23882896	23807520
24231716	24116576	8944291	22309474	20135799	20135799

批次統計

為與iftask/dinet相關的QPVC-DI效能開發了bulkstat方案。這對於從效能/負載角度監控DINET、服務埠和NPU利用率非常有用：

```
card schema iftask-dinet format EMS,IFTASKDINET,%date%,%time%,%dinet-rxpmts-curr%,%dinet-txpmts-curr%,%dinet-rxpmts-5minave%,%dinet-txpmts-5minave%,%dinet-rxpmts-15minave%,%dinet-txpmts-15minave%,%dinet-txdrops-curr%,%dinet-txdrops-5minave%,%dinet-txdrops-15minave%,%npuutil-now%  
file 2
```

```
port schema iftask-port format EMS,IFTASKPORT,%date%,%time%,%util-rxpmts-curr%,%util-txpmts-curr%,%util-rxpmts-5min%,%util-txpmts-5min%,%util-rxpmts-15min%,%util-txpmts-15min%,%util-txdrops-curr%,%util-txdrops-5min%,%util-txdrops-15min%  
file 3
```

```
card schema npu-util format EMS,NPUUTIL,%date%,%time%,%npuutil-now%,%npuutil-5minave%,%npuutil-15minave%,%npuutil-rxbytes-5secave%,%npuutil-txbytes-5secave%,%npuutil-rxbytes-5minave%,%npuutil-txbytes-5minave%,%npuutil-rxbytes-15minave%,%npuutil-txbytes-15minave%,%npuutil-rxpmts-5secave%,%npuutil-tpmts-5secave%,%npuutil-rxpmts-5minave%,%npuutil-tpmts-5minave%,%npuutil-rxpmts-15minave%,%npuutil-tpmts-15minave%
```