# The Battle Of The ZCs

### Who Is The Prettiest of Them All?



Jamal Hadi Salim | Nabil Bitar | Victor Nogueira | Pedro Tammela

Gemini: draw an image of two DRAM sticks playing a role in the theme of "mirror mirror on the wall, who is the prettiest of them all?"

### Work Done At Bloomberg

### Goals: RX+TX TCP ZC

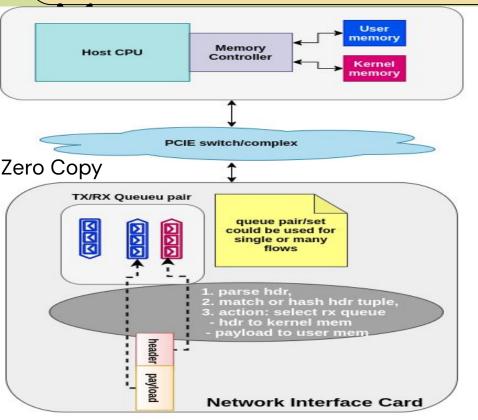
Investigate two new RX+TX TCP ZC techniques from an application perspective

- Baseline: traditional socket API app without any zero copy
- TCP Devmem and io\_uring ZC: development in motion at time of preparing this talk

<u>Note</u>: There are other established ZC techniques <u>we are not going to look at</u>

- Total Kernel bypass: DPDK and friends
- Semi-kernel bypass: AF\_XDP
- Sendfile (TX only)

## Zero Copy High Level



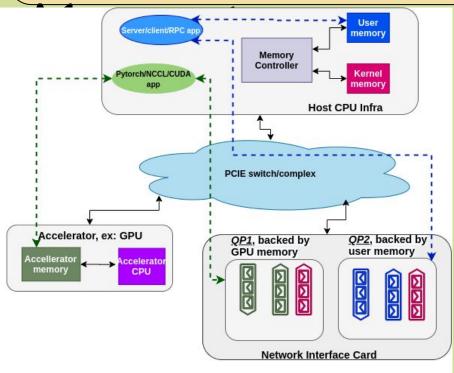
#### RX Control:

User memory(Blue) is bound to NIC buffers on rx of a queue pair

#### RX Fast path:

- 1. Incoming frame parsed for headers
- 2. header matched or hashed
- Action: Header placed onto kernel memory(red) and payload placed on user mapped memory(blue)
- Header goes over standard net stack and user app gets notified using either devmem or io\_uring

## Zero Copy High Level



User-space memory bound to NIC using udmabuf (Devmem)

#### io\_uring memory mapped ring

- Accelerator (e.g., GPU) memory bound to NIC using dmabuf
  - Devmem
  - Not currently supported by io\_uring



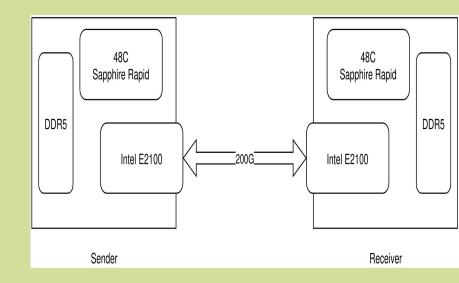
### Traffic Patterns Used

To quote David Ahern @netdevconf 0x18: *"The ultimate test is to fill ~<u>x00 gbps</u> with one flow"* 

- Asymmetric Client-Server
  - MTU large enough to cover MSS++
  - o IPv4
  - TCP Advertised mss: 4108B (4096B + 12B TCP options)
  - Very long flow analogous to Netperf TCP\_STREAM- single flow runs for 100s
  - 64MB app message
  - udmabuf used for devmem (both sides)

### Setup: CPU

- Motherboard: Intel Ruby Pass
- CPU: Single Intel(R) Sapphire Rapids CPU
  - Eagle Stream platform TDP 350W
  - 48C, Hyper-Threading off
- RAM: DDR5 128GB 4800 MT/s
- Application
  - We use our own home-brewed test program
  - CPU binding
    - Networking bound to 1 Core
    - User app bound to a different Core
  - Traffic: Single flow lands on queue O



### Setup: NIC

Intel IPU E2100:

(https://www.intel.com/content/www/us/en/products/details/network-io/ipu/adapter-e2100.html)

Used Single port connection back to back (2x100G Intel IPU E2100)

- using PCIE4x16 slot (should be able to handle 200Gbps)
- 2x100G with Port 0 at 200G config and Port 1 not in use
- Header split, SG, csum offload, TSO, HW GRO capability
- Ring size:
  - **RX:4096**
  - **TX:4096**

driver: idpf
version: 6.14.0-rc1+
firmware-version:
expansion-rom-version:
bus-info: 0000:6a:00.0
supports-statistics: yes
supports-test: no
supports-eeprom-access: no
supports-register-dump: no
supports-priv-flags: no



sysctl -w net.core.rmem\_max=536870912
sysctl -w net.core.wmem\_max=536870912
sysctl -w net.core.rmem\_default=16777216
sysctl -w net.core.wmem\_default=16777216
sysctl -w net.ipv4.tcp\_rmem="4096 87380 536870912"
sysctl -w net.ipv4.tcp\_wmem="4096 87380 536870912"

Receiver

sysctl -w net.core.optmem\_max=4194304

echo 16384 > /sys/devices/system/node/node0/hugepages/hugepages-2048kB/nr\_hugepages

sysctl -w vm.max\_map\_count=524240

sysctl –w net.core.optmem\_max=4194304

Sender

### General settings

- Kernel: 6.14.0\_rc1
- Extra patches for driver
  - Adapt idpf driver to devmem tcp (From Sridhar)
  - Fix issue in RSC flow (From Sridhar)
  - Enable BIG TCP
    - Driver change to 131072B
    - Kernel config: CONFIG\_MAX\_SKB\_FRAGS to 45
- io\_uring: zcrx v13 (from Pavel's repo)
- devmem: RFC V1 (December 2024 from Mina)
  - Bug fixes on devmem TX

### Setup: Measurement

- Interested in 3 metrics
  - Throughput, Power Consumption and CPU Usage
  - Results computed based on these 3 metrics
- Power consumption measured by an external device
  - Server as a 'black box'
  - Baseline power around 230 Watts
- Test duration: 100 seconds (> 1TB of data exchanged, so glitches even out)
  - Repeated 5 times, throw 1st and last and middle 3 averaged
- Governor: starts in power-save mode but ramps up within 100 seconds



### **Experiment Matrix**

- MTU Fixed: 4148B (4K MSS)
- Receiver side setup: permutation of SW GRO, HW GRO, BigTCP, no BigTCP
  - > Baseline (no ZC)
  - ≻ ZC
    - Devmem
    - io\_uring
- Sender side: BigTCP and no BigTCP
  - > Baseline (no ZC)
  - ≻ ZC
    - Devmem
    - io\_uring

### Here Be Dragons



Gemini prompt: " Draw an old map with dragons and other creatures "

### Some Bugs Were Fixed!

- Rewind the global binding tx\_iter by the dmabuf offset after tcp\_sendmsg\_locked finished
- Only return -EOPNOTSUPP if (!pool->dma\_map || !pool->dma\_sync) in page\_pool\_init (as it was before commit b400f4b8743)
- Add page pool stats to idpf driver

### io\_uring - iowait accounting

- io-uring trips the iowait cpu meta state gets worse with more CPUs
  - On RX 2 cpus this accounts for up to ~35% but for 1 cpu it accounts for ~5%
  - On TX (2 cpus) it accounts for ~87%
  - Pavel provided us a patch (back in 0x18) but it was not used in these tests

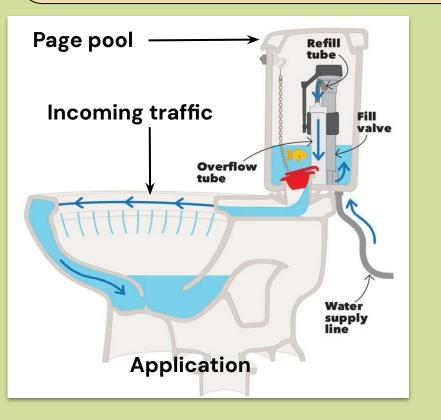
### IO\_URing - iowait accounting

CPU	POLL	C1	C1E	C6
0	0	0.01	0.01       57.06         6.68       41.16         5.41       43.19         1.23       58.13         4.23       40.04         0.01       51.67         9.43       27.83         3.16       40.48	
1	13.09	0.01       57.06         6.68       41.16         5.41       43.19         1.23       58.13         4.23       40.04         0.01       51.67         9.43       27.83         3.16       40.48         0       38.25		19.2
2	11.47	5.41	43.19	45.38
3	1.01	1.23	58.13	20.56
4	7.07	4.23	40.04	37.66
5	0	0.01	51.67	30.73
6	20.91	9.43	27.83	13.34
7	12.81	3.16 40.48		42.64
8	0.02	0	38.25	73.54
9	21.96	10.02	23.68	46.14
10	15.1	8.6	33.94	45.59
11	0	0.01	33.22	47.47
12	11.98	5.41	33.17	30.66
13	22.29	9.16	23.86	44.8
14	14.34	6.04	35.44	44.07
15	11.31	5.67	32.77	57.62

Improved idle states Reduced power consumption

CPU	POLL	C1	C1E	C6
0	0	0	0.58	67.21
1	0.47	1.56	1.4 42.03	
2	0.98	1.4 42.03		29.86
3	1.56	1.66	49.61	22.85
4	0.82	1.61	1.61 30.51	
5	0.02	0.04	0.04 19.82	
6	0	0 0.67		99.29
7	0	0	0.6	82.31
8	0.99	1.61		
9	0	0	0.32	83.25
10	0.06	0.11	25.89	52.4
11	0.47	1.59	40.61	53.88
12	0.9	1.72	46.9	44.29
13	1.11	1.16	42.77	42.01
14	0.79	1.67	42.32	52.22
15	0	0	0.58	89.64

### Getting The Plumbing Right



200Gbps @0.1ms RTT implies we need to fill a tank of ~2MB to cover BDP (Bandwidth delay product)

Assuming ~2MB inflight to the application, use 2xBDP (~4MB)

Fix RX ring: 4096 (per qp) x page size(4096B) = 16MB.

Compensate for app sitting on accepted data (48MB): overprovision to 64MB per qp for udmabuf

### Devmem RX: Receive Size

- We allocated Udmabuf size of 64MB
  - 4x the DMA RX size to factor in that the app may sit on the buffers for a while
- We observed varying recvmsg() read size also affects performance
  - Started with app read size of 512KB
  - In this scenario, buffers are not recycled fast enough, which triggers the page allocator more frequently (see "Bad")
- Read as much as 64MB improved our throughput by 30%
  - On average we are seeing ~875 tokens (recall each payload is 4KB)

#### Bad:

0verhead	Command	Shared Object	Symbol
9.40%	ksoftirqd/0	[kernel.kallsyms]	<pre>[k] gen_pool_alloc_algo_owner</pre>
8.52%	:33681	[kernel.kallsyms]	<pre>[k] tcp recvmsg dmabuf</pre>
8.11%	ksoftirqd/0	[kernel.kallsyms]	[k] gen pool has addr
7.98%	ksoftirqd/0	[kernel.kallsyms]	[k] gen pool free owner
3.24%	swapper	[kernel.kallsyms]	[k] gen pool alloc algo owner
- 3.11%	swapper	[idpf]	[k] idpf vport splitq napi poll
3.10%	swapper	[kernel.kallsyms]	[k] gen pool has addr
2.93%	swapper	[kernel.kallsyms]	[k] gen pool free owner
2.30%	:33681	[kernel.kallsyms]	[k] put cmsg

#### Good:

S	amples: 99	K of event	'cpu-cycles', Event count	(approx.): 90031202737
	0verhead	Command	Shared Object	Symbol
÷	31.21%	:22366	[kernel.kallsyms]	[k] tcp_recvmsg_dmabuf
÷	6.25%	:22366	[kernel.kallsyms]	[k] napi pp put page
÷	6.09%	:22366	[kernel.kallsyms]	[k] xas_store
÷	5.62%	:22366	[kernel.kallsyms]	[k] put_cmsg
÷	2.87%	:22366	[kernel.kallsyms]	<pre>[k] _raw_spin_lock_bh</pre>
÷	2.51%	:22366	[kernel.kallsyms]	[k] _raw_spin_lock
÷	2.19%	:22366	[kernel.kallsyms]	<pre>[k] xas find marked</pre>
÷	- 1.90%	:22366	[kernel.kallsyms]	[k] tcp_rcv_established

### io\_uring ring resizing

- In our tests we may have encountered cq events drops when testing on 1 CPU
  - This should not be possible, the default *IORING\_FEAT\_NODROP* is always on
  - Likely a bug, did not have time to investigate. Pavel gave us a tip:

In extreme memory starvation cases a CQE might get dropped, the syscall should eventually result in -EBADR<mark>.</mark> And can be checked with liburing ring->cq.koverflow.

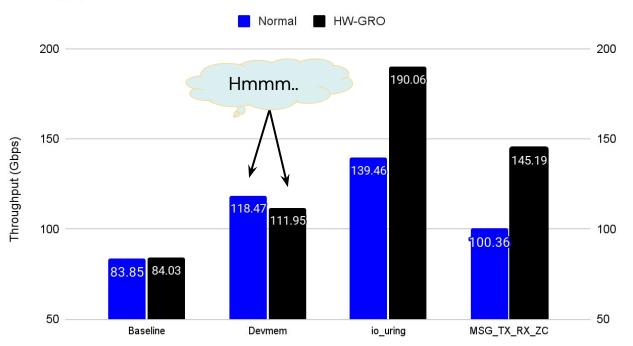
Note: Anytime we exhaust the page pool, the incoming data is wrapped in skbs and stored in socket queue then posted on CQ when space becomes available

This approach is taken by both io\_uring and devmem

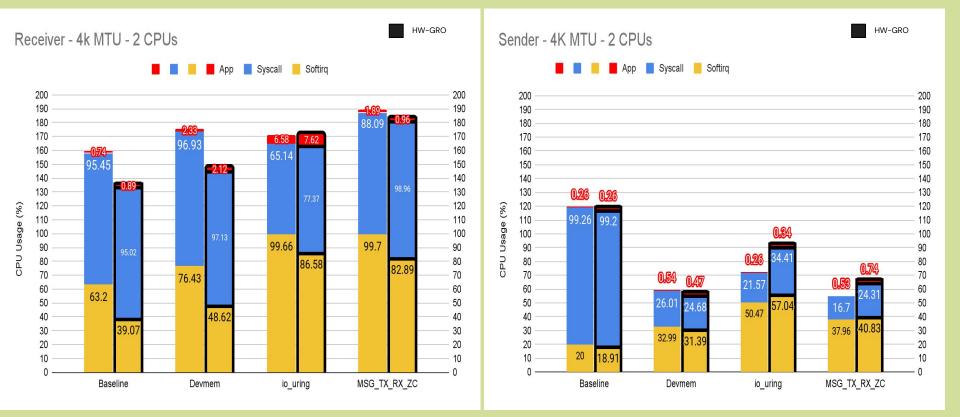
However best results are always achieved with proper provisioning

### First Impressions: Throughput

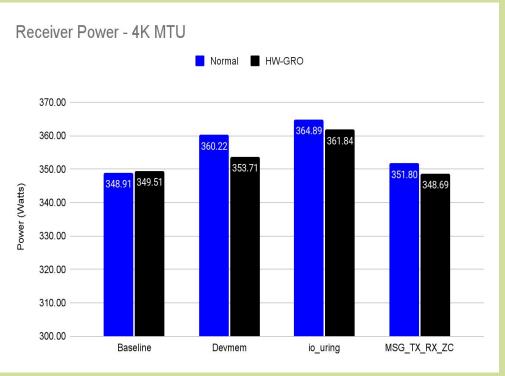
Throughput - 4k MTU - 2 CPUs

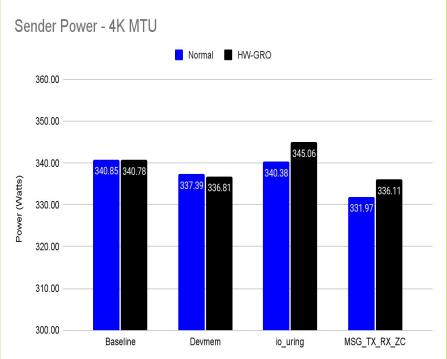






### Power





### Devmem vs io\_uring: S/W vs H/W GRO

	巴 Devmem - 2 CPUs - HW GRO (111.95 Gbps)
Samples: 199K of event 'cpu-cycles', Event count (approx.): 168906025843	Samples: 198K of event 'cpu-cycles', Event count (approx.): 142019659159
Overhead Command Shared Object Symbol	Overhead Command Shared Object Symbol
+ 18,61% :18102 [kernel.kallsyms] [k] tcp_recvmsq_dmabuf	+ 22,57% :18805 [kernel.kallsyms] [k] tcp_recvmsg_dmabuf
+ 8,54% swapper [idpf] [k] idpf_vport_splitg_napi_poll	+ 8,54% :18805 [kernel.kallsyms] [k] napi_pp_put_page
+ 6,29% :18102 [kernel.kallsyms] [k] napi_pp_put_page	+ 7,85% swapper [idpf] [k]idpf_vport_splitg_napi_poll
+ 5,71% swapper [kernel.kallsyms] [k] tcp_gro_receive	+ 4,31% swapper [libeth] [k] libeth_rx_recycle_slow
+ 5,21% swapper [kernel.kallsyms] [k] napi_pp_put_page	+ 4,12% :18805 [kernel.kallsyms] [k] put_cmsg
+ 3,63% :18102 [kernel.kallsyms] [k] put_cmsg	+ 3,87% swapper [kernel.kallsyms] [k] napi_pp_put_page
+ 2,93% :18102 [kernel.kallsyms] [k] xas_store	+ 3.41% :18805 [kernel.kallsyms] [k] xas_store
+ 2,12% :18102 [kernel.kallsyms] [k] xas find_marked	+ 2,64% swapper [kernel.kallsyms] [k] intel_idle
+ 1.82% :18102 [kernel.kallsyms] [k] xas_load	+ 2,37% :18805 [kernel.kallsyms] [k] xas_find_marked
+ 1,68% :18102 [kernel.kallsyms] [k] _raw_spin_lock_bh	+ 2,05% :18805 [kernel.kallsyms] [k] xas_load
+ 1,53% swapper [kernel.kallsyms] [k] napi_build_skb	+ 1.87% :18805 [kernel.kallsyms] [k] _raw_spin_lock_bh
+ 1,32% swapper [kernel.kallsyms] [k] tcp_v4_rcv	+ 1,46% :18805 [kernel.kallsyms] [k] tcp_rcv_established
+ 1,31% swapper [kernel.kallsyms] [k] dev_gro_receive	+ 1,37% swapper [kernel.kallsyms] [k] napi_build_skb
+ 1,26% swapper [kernel.kallsyms] [k] skb_release_data	+ 1,25% swapper [kernel.kallsyms] [k] skb_release_data
+ 1,26% swapper [kernel.kallsyms] [k] skb_release_data + 1,24% :18102 [kernel.kallsyms] [k] tcp_recvmsg_locked Devemom	+ 1.25% :18805 [kernel.kallsyms] [k] xas clear mark
+ 1,12% :18102 [kernel.kallsyms] [k] xas_clear_mark	+ 1,23% :18805 [kernel.kallsyms] [k] tcp_recvmsg_locked
+ 1,11% :18102 [kernel.kallsyms] [k] tcp_rcv_established	+ 1,12% :18805 [kernel.kallsyms] [k] skb_try_coalesce
+ 1,09% swapper [kernel.kallsyms] [k] skb_gro_receive	+ 1,07% swapper [kernel.kallsyms] [k] tcp v4 rcv
+ 0,87% swapper [kernel.kallsyms] [k]inet_lookup_established	+ 1,01% :18805 [kernel.kallsyms] [k] tcp_v4_do_rcv
+ 0,85% :18102 [kernel.kallsyms] [k] tcp_rcv_space_adjust	+ 0,79% swapper [kernel.kallsyms] [k]inet_lookup_established
+ 0,84% swapper [kernel.kallsyms] [k]napi_build_skb	+ 0.72% :18805 [kernel.kallsyms] [k] check preemption disabled
+ 0,82% swapper [kernel.kallsyms] [k] intel_idle	+ 0,69% :18805 [kernel.kallsyms] [k] sock_rfree
+ 0,82% :18102 [kernel.kallsyms] [k] skb_try_coalesce	+ 0,69% swapper [kernel.kallsyms] [k] net_rx_action
Tip: To browse sample contexts use perf reportsample 10 and select in context menu	Tip: To see call chains by final symbol taking CPU time (bottom up) use perf report -G
BP io_uring - 2 CPUs (139.46 Gbps)     Samples: 199K of event 'cou-cycles', Event count (approx.): 172695707526	io_uring - 2 CPUs - HW GRO (190.06 Gbps) Samples: 191K of event 'cou-cycles', Event count (approx.): 172013178812
⊕ <u>io_uring-2 CPUs (139.46 Gbps)</u> Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol	io_uring-2 CPUs-HW GRO(190.06 Gbps) Samples: 191K of event 'cpu-cycles', Event count (approx.): 1/20131/9812 Overhead Command Shared Object Symbol
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526	Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013179812
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol	Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013179812 Overhead Command Shared Object Symbol
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol * 8,235 :25567 [kernel.kallsyms] [k] to_zcrx_recv_skb * 6,66% ksofttrqd/0 [tdpf] [k] tdpf_vport_spittq_napi_poll * 6,66% ksofttrqd/0 [kernel.kallsyms] [k] to_grecetve	Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013179812 Overhead Command Shared Object Symbol + 10,74 :30719 [kernel.kallsyns] [k] to_zcrx_recv_skb + 4,47% swapper [kernel.kallsyns] [k] nogl_pp_put_page + 3,97% swapper [kdpf] [k] idpf_yport_spilt_napl_poll
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol = 0,259 :25507 [kernel.kallsyms] [k] to_zcrx_recv_skb = 0,468 ksoftrad/0 [idpf] [k] idpf_uport_split_napi_poll = 0,468 ksoftrad/0 [kernel.kallsyms] [k] top_aro_receive = 5,459 ksoftrad/0 [kernel.kallsyms]	Samples: 191K of event 'cpu-cycles', Event count (approx.): 120213170812       Overhead Command     Shared Object       Symples: 191K of event 'cpu-cycles', Event count (approx.): 120213170812       + 10,70%     :38719       + 4,47%     :swapper       - 4,47%     :swapper       - 10,70%     :38019       - 3,97%     :swapper       - 10,60%     :38719       - 10,60%     :38719       - 10,60%     :38719       - 10,60%     :38719
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol * 8,235 :25567 [kernet.kallsyms] [k] to_zcrx_recv_skb * 6,66% ksofttrqd/0 [kernet.kallsyms] [k] top_roport_spitta_napi_poll * 6,46% ksofttrqd/0 [kernet.kallsyms] [k] top_ropectypage * 5,45% ksofttrqd/0 [kernet.kallsyms] [k] ranj.pp_put_page * 4,06% :25567 [kernet.kallsyms] [k] ranj.pn.lock	Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013179812       Overhead Command     Shared Object       Symbol     + 10,10*       + 10,10*     :30719       [kernel.kallsyms]     [k] to_zcrx_recv_skb       + 4,47%     swapper       [kernel.kallsyms]     [k] napi_pp_put_page       + 3,67%     swapper       [kernel.kallsyms]     [k] tof_urot_spiltq_napi_poll       + 3,66%     :38719       [kernel.kallsyms]     [k] poli_lock       + 3,66%     :swapper       [k] row_spin_lock
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526       Overhead Command     Shared Object       * 0,255     :25507       * 6,456     ksoftriqd/0       * 1000000000000000000000000000000000000	Samples: 191K of event 'cpu-cycles', Event count (approx.): 12013170812       Overhead Command     Shared Object       Symples: 191K of event 'cpu-cycles', Event count (approx.): 120213170812       Overhead Command     Shared Object       Symples: 191K of event 'cpu-cycles', Event count (approx.): 120213170812       + 19,70%     :30719       + 3,97%     swapper       - 3,97%     swapper       - 4,47%     :307%       - 3,97%     swapper       - 4,17%     :101f_vport.splitde.napi_poll       + 3,61%     swapper       - Kernel.kallsyns]     [K] pol_lufe       + 3,61%     swapper       - Kernel.kallsyns]     [K] pol_lufe       + 2,90%     swapper       - Kernel.kallsyns]       - 2,00%     swapper       - Kernel.kallsyns       - 2,00%       - K       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       - 2,00%       -
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol * J.255 :22567 [kernel.kallsyms] [k] lo_zcrx_recv_skb * d,666 ksofttrqd/0 [kernel.kallsyms] [k] top_croct_splittq_napi_poll * d,868 ksofttrqd/0 [kernel.kallsyms] [k] napi_pp_put_page * d,964 :22567 [kernel.kallsyms] [k] rad_splin_lock * d,964 ksofttrqd/0 [kernel.kallsyms] [k] rad_splin_lock * d,964 ksofttrqd/0 [kernel.kallsyms] [k] tote_idle * 2,064 ksofttrqd/0 [kernel.kallsyms] [k] tote_idle	Samples: 191K of event 'cpu-cycles', Event count (approx.): 1720131720812       Overhead Command Shared Object Symbol       + 10, *C* :30719 [kernel.kallsyns] [k] lo_zcrx_recv_skb       + 4, *T% swapper [kernel.kallsyns] [k] npi_pp_put_page       + 3, 67% :30719 [kernel.kallsyns] [k] idpi_prot.spltt_napt_poll       + 3, 66% :30719 [kernel.kallsyns] [k] on_zcpi_lock       + 3, 66% :swapper [kernel.kallsyns] [k] opp_zc_alloc_netmens       + 2, 96% swapper [kernel.kallsyns] [k] idpi_prot_splt
Samples: 199K of event 'cpu-cycles', Event count (approx.): 1/2695707526       Overhead Command     Shared Object       9.259     (Evente, Lkallsyms]       16.100     Ksoftrigd/0       17.100     (K) (dof, Uprot, split, napi_poll       18.100     (K) (dof, Uprot, split, napi_poll       19.100     (K) (dof, Uprot, split, napi_poll       29.061     (Koftrigd/0       20.051     (Kortirdd/0       20.051     (Kortirdd/0       20.051     (Kortirdd/0       20.052     (K) (sbl_prot, split, all split, split, all split, split, all split, split, all split, all split, split, all sp	Samples: 191K of event 'cpu-cycles', Event count (approx.): 12013170812         Overhead Command       Shared Object         Symples: 191K of event 'cpu-cycles', Event count (approx.): 120213170812         Overhead Command       Shared Object         Symples: 191K of event 'cpu-cycles', Event count (approx.): 120213170812         + 19,70%       :30719         + 4,477       swapper [kernel.kallsyms] [k] napi_pput,page         + 3,67%       :3075         * 3,61%       :3092per         * 4,61%       :3071         * 3,61%       :3092per         * 10,70%       :2072_clast_nlckallsyms] [k] rav_spin.lck         * 3,61%       swapper         * (kernel.kallsyms] [k] poll_ufle         * 2,85%       swapper [kernel.kallsyms] [k] poll_ufle         * 2,25%       swapper [kernel.kallsyms] [k] jsb.retese_data         * 2,25%       swapper [kernel.kallsyms] [k] jsb.retese_data
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526         Overhead Command Shared Object Symbol         * 0.25% :25507       [kernel.kallsyms] [k] top.crx_recv_skb         * 0,46% ksoftirqd/0       [kernel.kallsyms] [k] top.crc_recv_skb         * 0,46% ksoftirqd/0       [kernel.kallsyms] [k] top.crc_recv_skb         * 0,46% ksoftirqd/0       [kernel.kallsyms] [k] top.crc_recves         * 3,45% ksoftirqd/0       [kernel.kallsyms] [k] top.grc_receve         * 3,45% ksoftirqd/0       [kernel.kallsyms] [k] top.grc_receve         * 3,22% swapper       [kernel.kallsyms] [k] top.grc_alloc_netmens         * 2,04% ksoftirqd/0       [kernel.kallsyms] [k] staf_relata         * 2,95% ksoftirqd/0       [kernel.kallsyms] [k] [k] staf_relase_data         * 2,95% ksoftirqd/0       [kernel.kallsyms] [k] [k] staf_receve	Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013170812       Overhead Command     Shared Object       Swnbol     Shared Object       Verhead Command     Sharelos Object       Verhead Swapper </th
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526         Overhead Command       Shared Object Symbol         * 0.259       :25307         (kernel.kallsyms]       [k] idpf.yort.splitq.napi_poll         * 0.469       ksoftirqd/0       [kernel.kallsyms]         * 0.469       ksoftirqd/0       [kernel.kallsyms]         * 0.469       ksoftirqd/0       [kernel.kallsyms]         * 0.469       ksoftirqd/0       [kernel.kallsyms]         * 1.495       xsoftirqd/0       [kernel.kallsyms]         * 1.925       support       [k] napi_point_point_page         * 4.945       :25307       [kernel.kallsyms]         * 1.925       support       [k] intel_idle         * 2.964       ksoftirqd/0       [kernel.kallsyms]         * 2.964       ksoftirqd/0       [kernel.kallsyms]         * 1.955       ksoftirqd/0       [kernel.kallsyms]         * 1.945       ksoftirqd/0<	Samples: 191K of event 'cpu-cycles', Event count (approx.): 17201370812         Overhead Command       Shared Object         Symples: 191K of event 'cpu-cycles', Event count (approx.): 17201370812         Overhead Command       Shared Object         Symples: 191K of event 'cpu-cycles', Event count (approx.): 17201370812         + 1, 976       :38719         (kernel.kallsyms)       [K] napi_pp_nut_page         + 3, 67%       :38719         [kernel.kallsyms]       [K] rout_spin.lck         + 3, 61%       swapper         [kernel.kallsyms]       [K] rout_spin.lck         + 3, 61%       swapper         [kernel.kallsyms]       [K] plu_blu_cle         + 2, 28%       swapper         [kernel.kallsyms]       [K] lobeth_rx_recycle_slow         + 2, 2,85%       swapper         [kernel.kallsyms]       [K] sks_release_data         + 2, 2,85%       swapper         [kernel.kallsyms]       [K] net_rx_recycle_slow         + 2, 2,85%       swapper         [kernel.kallsyms]       [K] plage_pool_alloc_frag_netmen         + 2,98%       swapper         [kernel.kallsyms]       [K] net_rx_recycle_slow         + 2,98%       swapper         [kernel.kallsymsi       [K] net_rx_recycle_slow
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526         Overhead Command Shared Object Symbol         * 2,25%       22567         [k] for event 'cpu-cycles', Event count (approx.): 172695707526         * 6,26%       22567         [k] for event 'cpu-cycles', Event count (approx.): 172695707526         * 6,26%       22567         [k] top 'cpc-cycles', ecv.ysb         * 6,46%       ksoftirqd/0         [k] top 'cpc-cectve         * 5,45%       ksoftirqd/0         [kernel.kallsyms]       [k] top 'cpc-cectve         * 6,46%       ksoftirqd/0         [kernel.kallsyms]       [k] rev.pinlock         * 3,22%       supper         [kernel.kallsyms]       [k] top 'cpc-celve         * 2,04%       ksoftirqd/0         [kernel.kallsyms]       [k] top 'cpc-celve         * 2,04%       ksoftirqd/0         [kernel.kallsyms]       [k] bop 'cpc-calloc_netmens         * 2,39%       ksoftirqd/0         [kernel.kallsyms]       [k] bop 'cpc-ceclve         * 1,84%       ksoftirqd/0         [kernel.kallsyms]       [k] bop 'cpc-ceclve         * 1,84%       ksoftirqd/0         [kernel.kallsyms]       [k] bop 'cpc-ceclve         * 1,84%	Samples: 191K of event 'cpu-cycles', Event count (approx.): 12013170812       Overhead Command Shared Object Symbol       + 10,78: :30719       Kernel.kallsyms]       K1 log_crx_recv_skb       + 4,47% swapper       [Kernel.kallsyms]       [K] log_crx_recv_skb       + 3,67% swapper       [Kernel.kallsyms]       [K] log_crx_recv_skb       + 3,66% :38719       [Kernel.kallsyms]       [K] log_crx_recv_skb       + 2,06% swapper       [Kernel.kallsyms]       [K] log_crx_recv_slm       * 2,06% swapper       [Kernel.kallsyms]       [K] log_crx_recv_slm       * 2,20% swapper       [Kernel.kallsyms]       [K] log_crx_recv_slm       * 2,20% swapper       [Kernel.kallsyms]       [K] stb_release_data       * 2,21% swapper       [Kernel.kallsyms]       [K] page_cool_alloc_frag_netmen       * 1,94% swapper       [Kernel.kallsyms]       [K] loff_trag_cleand       * 1,94% swapper       [Kernel.kallsyms]       [K] loff_trag_cleand       * 1,94% swapper       [K] loff_trag_cleand       * 1,94% swapper       [K] loff_trag_cleand       * 1,94% swapper       [K] loff_trag_leland       * 1,96% swapper       [K]
Samples: 199K of event 'cup-cycles', Event count (approx.): 172695707526         Overhead Command       Shared Object Symbol         * 0,255       :25507         (kernel.kallsyms]       [k] idpf.yorx.pitt_napi_poll         * 0,465       ksoftirqd/0         Kernel.kallsyms]       [k] idpf.yorp.put_page         * 0,465       ksoftirqd/0         Kernel.kallsyms]       [k] idpf.yorp.put_page         * 4,045       :25507         Kernel.kallsyms]       [k] intp.yont_page         * 4,045       :25507         Kernel.kallsyms]       [k] intp.intp.kok         * 3,225       supp.t         * 4,045       :25507         Kernel.kallsyms]       [k] intp.intp.kok         * 2,045       ksoftirqd/0       [kernel.kallsyms]         * 2,345       ksoftirqd/0       [kernel.kallsyms]         * 1,855       ksoftirqd/0       [kernel.kallsyms]         * 1,857       ksoftirqd/0       [kernel.kallsyms]         * 1,785       <	<pre>Samples: 391K of event 'cpu-cycles', Event count (approx.): 120013170812 Overhead Command Shared Object Symph + 19,70% :30719 [kernel.kallsyms] [k] loc_2crx_recv_skb + 4,47% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,67% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,65% swapper [kernel.kallsyms] [k] nopi_pot_inleck + 3,61% swapper [kernel.kallsyms] [k] nopi_pot_inleck + 2,85% swapper [kernel.kallsyms] [k] join_z.clinc.netmens + 2,25% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 2,85% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 2,85% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 2,95% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 1,94% swapper [kernel.kallsyms] [k] not_rx_clean.compla + 1,75% swapper [kernel.kallsyms] [k] loft_rx_clean.compla + 1,75% swapper [kernel.kallsyms] [k] jor_slaw.compla + 1,95% swapper [kernel.kallsyms] [k] jor_slaw.compla</pre>
<pre>Samples: 199K of event 'cpu-cpcles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol 4 2,225: 125507 [kernel.kallsyms] [k] cp_crx_recv_skb 4 0,006: ksoftirqd/0 [kernel.kallsyms] [k] tp_pro_recetve 5 ,458: ksoftirqd/0 [kernel.kallsyms] [k] rop_put_page 4 ,004: 125507 [kernel.kallsyms] [k] rop_put_page 4 ,004: ksoftirqd/0 [kernel.kallsyms] [k] rop_ccalloc_netmens 8 ,229: ksoftirqd/0 [kernel.kallsyms] [k] kop_cc_alloc_netmens 4 , 2,39% ksoftirqd/0 [kernel.kallsyms] [k] kop_cc_alloc_netmens 4 , 2,39% ksoftirqd/0 [kernel.kallsyms] [k] kb_release_data 4 , 4,4% ksoftirqd/0 [kernel.kallsyms] [k] shp_receetve 4 , 1,84% ksoftirqd/0 [kernel.kallsyms] [k] shp_receetve 4 , 1,84% ksoftirqd/0 [kernel.kallsyms] [k] shp_receetve 4 , 1,84% ksoftirqd/0 [kernel.kallsyms] [k] shp_tbuild_skb 4 , 1,7% ksoftirqd/0 [kernel.kallsyms] [k] shp_ge_receive 4 , 1,7% ksoftirqd/0 [kernel.kallsyms] [k] shp_ge_receive 5 , 1,7% ksoftir</pre>	Samples: 191K of event 'cpu-cycles', Event count (approx.): 12013120812         Overhead Command Shared Object Symbol         + 1,78: :39719         + 4,47%, swapper [kernel.kallsyms] [k] napi_pp_nut_psge         + 4,47%, swapper [kernel.kallsyms] [k] napi_pp_nut_psge         + 3,67%, swapper [kernel.kallsyms] [k] napi_pp_nut_psge         + 3,66% :38719         + kernel.kallsyms] [k] obj_pz_calloc_tetmens         + 2,80%, swapper [kernel.kallsyms] [k] boj_pz_calloc_tetmens         + 2,20%, swapper [kernel.kallsyms] [k] boj_pz_calloc_trag.netmens         + 2,212       swapper [kernel.kallsyms] [k] page_pool_alloc_frag.netmen         + 1,04%, swapper [kernel.kallsyms] [k] hoj_callso_frag.netmen         + 1,04%, swapper [kernel.kallsyms] [k] lubpt_rx_clean_complq         + 1,04%, swapper [kernel.kallsyms] [k] lubpt_rx_clean_complq         + 1,06%, swapper [kernel.kallsyms] [k] lubpt_rx_lean_lean_lean_lean_lean_lean_lean_lean
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526         Overhead Command       Shared Object         9,25:       :25507         (kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         6,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,45%       :25507         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 1,25%       skoftirqd/0       [kernel.kallsyms]         * 2,36%       ksoftirqd/0       [kernel.kallsyms]         * 1,35%       ksoftirqd/0       [kernel.kallsyms]         * 1,35%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kerne	<pre>Samples: 391K of event 'cpu-cycles', Event count (approx.): 120013170812 Overhead Command Shared Object Symph + 19,70% :30719 [kernel.kallsyms] [k] loc_zcrx_recv.sbb + 4,475, swapper [kernel.kallsyms] [k] nopi_pput_page + 3,67% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,61% swapper [kernel.kallsyms] [k] raw.spin.lck + 3,61% swapper [kernel.kallsyms] [k] to_pz_calloc_netmens + 2,25% swapper [kernel.kallsyms] [k] shortelese.gdata + 2,85% swapper [kernel.kallsyms] [k] shortelese.gdata + 2,85% swapper [kernel.kallsyms] [k] shortelese.gdata + 2,28% swapper [kernel.kallsyms] [k] shortelese.gdata + 2,28% swapper [kernel.kallsyms] [k] shortelese.gdata + 1,94% swapper [kernel.kallsyms] [k] not_rx_recycle.slow + 1,94% swapper [kernel.kallsyms] [k] not_rx_clean.comptq + 1,95% swapper [kernel.kallsyms] [k] idpr_tx_clean.comptq + 1,62% swapper [kernel.kallsyms] [k] idpr_tx_clean.comptq + 1,62% swapper [kernel.kallsyms] [k] idr_tt_clean.comptq + 1,64% swapper [kernel.kallsyms] [k] idr_tt_cl</pre>
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172695707526         Overhead Command       Shared Object         9,25:       :25507         (kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         6,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,46%       Ksoftirqd/0         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 0,45%       :25507         Kernel.kallsyms)       [k] idpf.yport.splitt_napi_poll         * 1,25%       skoftirqd/0       [kernel.kallsyms]         * 2,36%       ksoftirqd/0       [kernel.kallsyms]         * 1,35%       ksoftirqd/0       [kernel.kallsyms]         * 1,35%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kernel.kallsyms]         * 1,37%       ksoftirqd/0       [kerne	<pre>Samples: 391K of event 'cpu-cycles', Event count (approx.): 3/2013/2012 Overhead Command Shared Object Symbol + 19,70% :30719 [kernel.kallsyms] [k] noz_crx_recv_skb + 4,47% swapper [kernel.kallsyms] [k] noz_pp_ut_page + 3,97% swapper [kernel.kallsyms] [k] row_spin_lock + 3,01% swapper [kernel.kallsyms] [k] row_spin_lock + 3,01% swapper [kernel.kallsyms] [k] poll_idle + 2,00% swapper [kernel.kallsyms] [k] opz_alloc_netmens + 2,25% swapper [kernel.kallsyms] [k] skp_release_data + 2,85% swapper [kernel.kallsyms] [k] pol_idle + 3,61% swapper [kernel.kallsyms] [k] pol_idle + 2,26% swapper [kernel.kallsyms] [k] skp_release_data + 2,26% swapper [kernel.kallsyms] [k] pol_idle_rag.netmen + 1,29% swapper [kernel.kallsyms] [k] pol_idle_rag.netmen + 1,29% swapper [kernel.kallsyms] [k] row_spin.lock + 1,69% swapper [kernel.kallsyms] [k] row_spin.lock + 1,26% swapper [kernel.kallsyms] [k] row_spin.lock + 1,55% swapper [kernel.kallsyms] [k] row_spin.lock + 1,54% swapper [kernel.kallsyms] [k] row_spin.lock + 1,54% swapper [kernel.kallsyms] [k] row_spin.lock + 1,54% swapper [kernel.kallsyms] [k] [k] [k] [k] [k] [k] [k] [k] [k] [k</pre>
Samples: 199K of event 'cup-cycles', Event count (approx.): 172695707526         Overhead Command       Shared Object         9,25:       :25967         1, error, 1, and Command       Shared Object         9,25:       :25967         1, error, 1, and Command       Shared Object         1, error, 1, and Command       Kardtragd/0         1, error, 1, and Command       Kardtragd/0         1, error, 1, and Command       Kardtragd/0         1, error, 1, and 1,	<pre>Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013170812 Overhead Command Shared Object Symph + 19,70% :30719 [kernel.kallsyms] [k] loc_zcrx_recv.sbb + 4,475, swapper [kernel.kallsyms] [k] nopi_pput_page + 3,67% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,65% swapper [kernel.kallsyms] [k] raw.spin.lck + 3,61% swapper [kernel.kallsyms] [k] op_z.calloc.netmens + 2,25% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 2,25% swapper [kernel.kallsyms] [k] jobt_calloc.netmens + 2,28% swapper [kernel.kallsyms] [k] jobt_calloc.netmens + 2,28% swapper [kernel.kallsyms] [k] jobt_calloc.netmens + 1,94% swapper [kernel.kallsyms] [k] jobt_calloc.netmens + 1,95% swapper [kernel.kallsyms] [k] jobt_calloc.netmen + 1,95% swapper [kernel.kallsyms] [k] jobt_calloc.frag.netmen + 1,95%</pre>
<pre>Samples: 199K of event 'cpu-cycles', Event count (approx.): 172095707520 Overhead Command Shared Object Symbol</pre>	<pre>Samples: 391K of event 'cpu-cycles', Event count (approx.): 3/2013/2012 Overhead Command Shared Object Symbol + 19,70% :30719 [kernel.kallsyms] [k] not_cpu-put_page + 3,97% swapper [kernel.kallsyms] [k] not_cpu-put_page + 3,97% swapper [kernel.kallsyms] [k] rout_splttg.napi_poll + 3,06% swapper [kernel.kallsyms] [k] rout_splttg.napi_cpul + 3,01% swapper [kernel.kallsyms] [k] rout_splttg.napi_cpul + 2,05% swapper [kernel.kallsyms] [k] to_pt_z_alloc.netmens + 2,25% swapper [kernel.kallsyms] [k] skp_relesse_data + 2,06% swapper [kernel.kallsyms] [k] page_pool_alloc_frag.netmen + 1,96% swapper [kernel.kallsyms] [k] not_spl_compt + 1,66% swapper [kernel.kallsyms] [k] not_spl_compt + 1,66% swapper [kernel.kallsyms] [k] not_spl_compt + 1,65% swapper [kernel.kallsyms] [k] not_stablshed + 1,</pre>
Samples: 199K of event 'cup-cycles', Event count (approx.): 172695707526         Overhead Command       Shared Object         Symbol       8,253         • 0,460       Ksoftriq470         • 6,460       Ksoftriq470         • 6,470       Ksoftriq470         • 7,375       Ksoftriq470         • 1,375       Ksoftriq470         • 1,375       Ksoftriq470         • 1,375       Ksoftriq470         • 1,475       Ksoftriq470         • 1,375       Ksoftriq470         • 1,375       Ksoftriq470         • 1,375       Ksoftriq470         • 1,475       Ksoftriq470         • 1,476       Ksoftriq470	<pre>Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013170812 Overhead Command Shared Object Symph + 10,70% :30719 [kernel.kallsyms] [k] loc_zcrx_recv_sbb + 4,47% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,67% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,65% swapper [kernel.kallsyms] [k] raw_spin.lck + 2,05% swapper [kernel.kallsyms] [k] nopi_zcalloc.netmens + 2,25% swapper [kernel.kallsyms] [k] locp_zcalloc.netmens + 2,25% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 2,28% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 2,28% swapper [kernel.kallsyms] [k] lobth_rx_recycle_slaw + 1,94% swapper [kernel.kallsyms] [k] not_rx_scalloc.netmens + 1,95% swapper [kernel.kallsyms] [k] not_rx_scalloc.frag.netmen + 1,95% swapper [kernel.kallsyms] [k] int_idle_train_oppi_q + 1,65% swapper [kernel.kallsyms] [k] int_idle_train_lock + 1,62% swapper [kernel.kallsyms] [k] int_idle_train_lock + 1,65% swapper</pre>
<pre>Samples: 199K of event 'cpu-cycles', Event count (approx.): 172095707520 Overhead Command Shared Object Symbol</pre>	<pre>Samples: 391K of event 'cpu-cycles', Event count (approx.): 372013720812 Overhead Command Shared Object Symbol + 19,70% :30719 [kernel.kallsyms] [k] nop.pp.pt.page + 3,97% swapper [kernel.kallsyms] [k] nopt.pp.pt.page + 3,97% swapper [kernel.kallsyms] [k] rout.pput.page + 3,97% swapper [kernel.kallsyms] [k] rout.pput.page + 3,01% swapper [kernel.kallsyms] [k] rout.pput.page + 2,05% swapper [kernel.kallsyms] [k] rout.pput.page + 2,25% swapper [kernel.kallsyms] [k] sbc.plesse_data + 2,26% swapper [kernel.kallsyms] [k] sbc.plesse_data + 2,26% swapper [kernel.kallsyms] [k] page_pool_alloc_frag.netmen + 1,96% swapper [kernel.kallsyms] [k] net_rx_action + 1,96% swapper [kernel.kallsyms] [k] net_rx_action + 1,96% swapper [kernel.kallsyms] [k] net_rk_alten complq + 1,96% swapper [kernel.kallsyms] [k]teilowup.stablished + 1,27% swapper [kernel.kallsyms] [k] top.cv_ercw + 0,96% :39719 [kernel.kallsyms] [k] sock_rfree</pre>
Samples: 199K of event 'cup-cycles', Event count (approx.): 172695707526 Overhead Command Shared Object Symbol 0,253: 22507 [kernel.kallsyms] [k] cl., 2crx, recv_skb 0,468 ksoftirqd/0 [idf] [k] idf/1prot.split_napi_poll 0,468 ksoftirqd/0 [kernel.kallsyms] [k] top gro_receive 0,468 ksoftirqd/0 [kernel.kallsyms] [k] cl. api_pout_page 4,045: 22507 [kernel.kallsyms] [k] cl. api_pout_page 2,268 ksoftirqd/0 [kernel.kallsyms] [k] idf. jop.c. alloc. netwens 2,268 ksoftirqd/0 [kernel.kallsyms] [k] sb. pelease.data 1,858 ksoftirqd/0 [kernel.kallsyms] [k] gro_por_ceceive 1,858 ksoftirqd/0 [kernel.kallsyms] [k] gro_por_ceceive 1,858 ksoftirqd/0 [kernel.kallsyms] [k] sb. pelease.data 1,858 ksoftirqd/0 [kernel.kallsyms] [k] sb. gro_receive 1,778 ksoftirqd/0 [kernel.kallsyms] [k] gro_por_celive 1,458 ksoftirqd/0 [kernel.kallsyms] [k] idf_tr_clean_complq 1,458 ksoftirqd/0 [kernel.kallsyms] [k] idf_tr_clean_complq 1,615: 25577 [kernel.kallsyms] [k] cl. gr_creceive 1,615: ksoftirqd/0 [kernel.kallsyms] [k] cl. gr_creceive 1,615: ksoftirqd/0 [kernel.kallsyms] [k] int_gr_creceive 1,615: ksoftirqd/0 [ker	<pre>Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013170812 Overhead Command Shared Object Symph. + 10,709: 130719 [kernel.kallsyms] [k] loc_zcrx_recv_skb + 4,475 swapper [kernel.kallsyms] [k] nopi_pput_page + 3,575 swapper [kernel.kallsyms] [k] nopi_pput_page + 3,575 swapper [kernel.kallsyms] [k] raw_spin.lck + 2,305 swapper [kernel.kallsyms] [k] nopi_zc_alloc.netmens + 2,255 swapper [kernel.kallsyms] [k] sopi_zcalloc.netmens + 2,255 swapper [kernel.kallsyms] [k] sopi_zcalloc.netmens + 2,285 swapper [kernel.kallsyms] [k] sopi_zcalloc.netmens + 2,285 swapper [kernel.kallsyms] [k] sopi_zcalloc.netmens + 1,945 swapper [kernel.kallsyms] [k] sopi_zcalloc.netmens + 1,945 swapper [kernel.kallsyms] [k] idpf_tx.clean.complq + 1,945 swapper [kernel.kallsyms] [k] idpf_tx.clean.complq + 1,955 swapper [kernel.kallsyms] [k] intel_idle_xstat + 1,405 swapper [kernel.kallsyms] [k] intel_idle_xstat + 4,405 swapper [kernel.kallsyms]</pre>
<pre>Samples: 199K of event 'cpu-cycles', Event count (approx.): 172095707520 Overhead Command Shared Object Symbol • 0,255 :23507 [kernel.kallsyms] [k] top_gro_receive • 0,465 ksoftirqd/0 [kernel.kallsyms] [k] top_gro_receive • 0,468 ksoftirqd/0 [kernel.kallsyms] [k] top_gro_receive • 4,844 :22507 [kernel.kallsyms] [k] top_gro_receive • 2,644 ksoftirqd/0 [kernel.kallsyms] [k] top_gro_receive • 2,644 ksoftirqd/0 [kernel.kallsyms] [k] top_gro_receive • 2,644 ksoftirqd/0 [kernel.kallsyms] [k] top_gro_ceive • 2,644 ksoftirqd/0 [kernel.kallsyms] [k] top_gro_receive • 1,658 ksoftirqd/0 [kernel.kallsyms] [k] bg pg_co_alloc_netmens • 1,658 ksoftirqd/0 [kernel.kallsyms] [k] bg pg_co_alloc_rrag_netmem • 1,758 ksoftirqd/0 [kernel.kallsyms] [k] bg pg_co_alloc_arg_netmem • 1,758 ksoftirqd/0 [kernel.kallsyms] [k] bg pg_co_alloc_rrag_netmem • 1,758 ksoftirqd/0 [kernel.kallsyms] [k] cherk_pro_receive • 1,658 ksoftirqd/0 [kernel.kallsyms] [k] cherk_pro_receive • 1,758 ksoftirqd/0 [kernel.kallsyms] [k] cherk_pro_receive • 1,958 ksoftirqd/0 [kernel.kallsyms] [k] cherk_pro_receive • 1,958 ksoftirqd/0 [kernel.kallsyms] [k] cherk_pro_receive • 0,958 ksoftirqd/0 [kernel.kallsyms] [k] top_rcv_established • 0,958 ksoftirqd/0 [kernel.kallsyms] [k] top_rcv_established • 0,958 ksoftirqd/0 [kernel.kallsyms] [k] top_rcv_established • 0,958 ksoftirqd/0 [kernel.kallsyms] [k] ray_rcv</pre>	<pre>Samples: 39% of event 'cpu-cycles', Event count (approx.): 172013170812 Overhead Command Shared Object Symbol + 10,70% :30719 [kernel.kallsyms] [k] nop.pp.pt.page + 3,97% swapper [kernel.kallsyms] [k] nopt.pp.pt.page + 3,97% swapper [kernel.kallsyms] [k] row.spin.lock + 3,01% swapper [kernel.kallsyms] [k] row.spin.lock + 3,01% swapper [kernel.kallsyms] [k] row.spin.lock + 2,00% swapper [kernel.kallsyms] [k] sob_rlesse_data + 2,25% swapper [kernel.kallsyms] [k] sob_rlesse_data + 2,06% swapper [kernel.kallsyms] [k] page_pool_alloc_frag.netmen + 1,94% swapper [kernel.kallsyms] [k] not_rx_recycle_slow + 1,94% swapper [kernel.kallsyms] [k] not_rx_recycle_stablished + 1,40% swapper [kernel.kallsyms] [k] not_ry_recycle_stablished + 1,22% swapper [kernel.kallsyms] [k] not_ry_recycle_stablished + 1,23% swapper [kernel.kallsyms] [k] sock_rfree + 0,96% :38719 [kernel.</pre>
Samples: 199K of event 'cpu-cycles', Event count (approx.): 172095707520 Overhead Command Shared Object Symbol * 0,250; 125567 [kernel.kallsyms] [k] to_zcrx_recv_skb * 0,400; ksoftirad/0 [tdyf1] [k] tdyf_uprot_splitq_napi_poll * 0,400; ksoftirad/0 [kernel.kallsyms] [k] to_gro_recetve * 0,404; 125567 [kernel.kallsyms] [k] to_gro_recetve * 0,404; 125567 [kernel.kallsyms] [k] to_gro_recetve * 2,404; ksoftirad/0 [kernel.kallsyms] [k] to_gro_recetve * 2,404; ksoftirad/0 [kernel.kallsyms] [k] to_gro_recetve * 1,804; ksoftirad/0 [kernel.kallsyms] [k] sbr_elease.data * 1,805; ksoftirad/0 [kernel.kallsyms] [k] sbr_elease.data * 1,804; ksoftirad/0 [kernel.kallsyms] [k] sbr_ereaption_alloc_frag_netmem * 1,494; ksoftirad/0 [kernel.kallsyms] [k] sbr_elease.data * 1,494; ksoftirad/0 [kernel.kallsyms] [k] sbr_ereaption_alloc_frag_netmem * 1,494; ksoftirad/0 [kernel.kallsyms] [k] to_gro_receive * 1,614; ksoftirad/0 [kernel.kallsyms] [k] to_gro_receive * 0,994; kso	<pre>Samples: 191K of event 'cpu-cycles', Event count (approx.): 172013170812 Overhead Command Shared Object Symph + 10,70% :30719 [kernel.kallsyms] [k] loc_zcrx_recv_sk0 + 3,67% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,67% swapper [kernel.kallsyms] [k] nopi_pput_page + 3,65% swapper [kernel.kallsyms] [k] nopi_pput_page + 2,05% swapper [kernel.kallsyms] [k] nopi_zput_alloc.netmess + 2,25% swapper [kernel.kallsyms] [k] sopi_zcalloc.netmess + 2,25% swapper [kernel.kallsyms] [k] sopi_zcalloc.netmess + 2,28% swapper [kernel.kallsyms] [k] sopi_zcalloc.netmess + 2,28% swapper [kernel.kallsyms] [k] sopi_zcalloc.netmess + 2,28% swapper [kernel.kallsyms] [k] sopi_zcalloc.netmess + 1,94% swapper [kernel.kallsyms] [k] sopi_zcalloc.netmes + 1,94% swapper [kernel.kallsyms] [k] not_rx_recycle_slaw + 1,95% swapper [kernel.kallsyms] [k] not_rx_ristored + 1,95% swapper [kernel.kallsyms] [k] not_rx_ristored + 1,95% swapper [kernel.kallsyms] [k] not_slaws + 1,95% swapper [kernel.kallsyms] [k] intel_idie + 1,40% swapper [kernel.kallsyms] [k] intel_idie_xstat + 4,40% (swapper [kernel.kallsyms] [k] inte</pre>

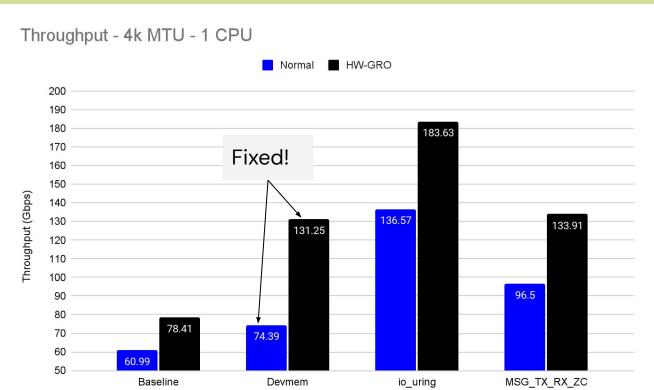
### Syscall Costs

:14342 (14342), 4745666 events, 100.0%

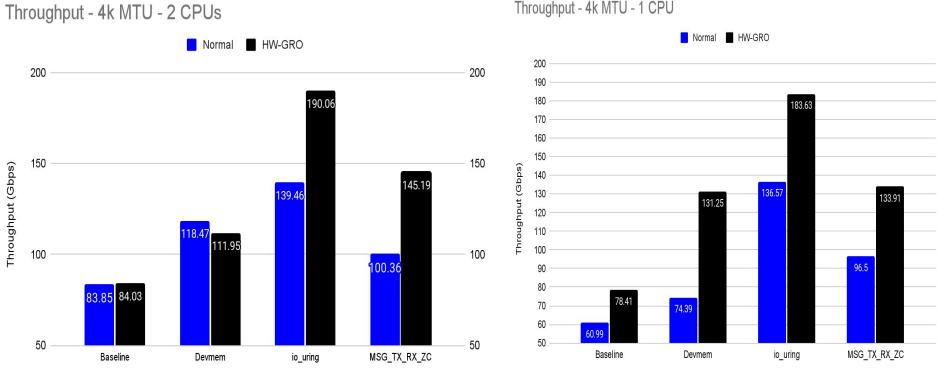
#### Devmem: 2 CPUs

syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
recvmsg setsockopt	615359 530116			0.001 0.001 0.001	0.112 0.043	0.371 0.284	 0.03% 0.03%
ioctl	1224901	0	1232.993	0.001	0.001	0.604	0.09%
epoll_wait	1835	0	628.492	0.000	0.343	100.162	28.36%
tcpzc2 (10500), 2	21032201 ev	vents, 10	00.0%			io_urii	ng: 2 CPUs
syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
io uring enter	10519910	0	76548.951	0.000	0.007	0.450	0.02%

### Throughput (I CPU)



### Throughput (1 vs 2 CPU)



Throughput - 4k MTU - 1 CPU

### Perf Devmem GRO (1 vs 2 CPUs)

2	Devmem - 2 CPUs (118.47 Gbps)	日本 Devmem - 2 CPUs - HW GRO (111.95 Gbps)											
	mples: 199K of event 'cpu-cycles', Event count (approx.): 168906025843					Samples: 198K of event 'cpu-cycles', Event count (approx.): 142019659159							
Overhead Command	Shared Object Symbol	0v	verhead	Comma	nand Shared	d Object !	Symbol						
+ 18,61% :18102	[kernel.kallsyms] [k] tcp_recvmsg_dmabuf	+		:1880			[k] tcp_recvmsg_dmabuf						
+ 8,54% swapper	[idpf] [k] idpf_vport_splitq_napi_poll	+		:1880	305 [kerne	el.kallsyms]	[k] napi_pp_put_page						
+ 6,29% :18102	[kernel.kallsyms] [k] napi_pp_put_page	+		swapp			[k] idpf_vport_splitq_napi_poll						
+ 5,71% swapper	[kernel.kallsyms] [k] tcp_gro_receive	+					[k] libeth_rx_recycle_slow						
+ 5,21% swapper	[kernel.kallsyms] [k] napi_pp_put_page	+				el.kallsyms]							
+ 3,63% :18102	[kernel.kallsyms] [k] put_cmsg	+					[k] napi_pp_put_page						
+ 2,93% :18102	[kernel.kallsyms] [k] xas_store	+				el.kallsyms]							
+ 2,12% :18102	[kernel.kallsyms] [k] xas_find_marked	+					[k] intel_idle						
+ 1,82% :18102	[kernel.kallsyms] [k] xas_load	+					[k] xas_find_marked						
+ 1,68% :18102	[kernel.kallsyms] [k] _raw_spin_lock_bh	. +				el.kallsyms]							
+ 1,53% swapper	[kernel.kallsyms] [k] napi_build_skb	+					[k] _raw_spin_lock_bh						
+ 1,32% swapper	[kernel.kallsyms] [k] tcp_v4_rcv	+					[k] tcp_rcv_established						
+ 1,31% swapper	[kernel.kallsyms] [k] dev_gro_receive	+					[k] napi_build_skb						
+ 1,26% swapper	[kernel.kallsyms] [k] skb_release_data						[k] skb_release_data						
+ 1,24% :18102	[kernel.kallsyms] [k] tcp_recvmsg_locked	+					[k] xas_clear_mark						
+ 1,12% :18102	[kernel.kallsyms] [k] xas_clear_mark	+					[k] tcp_recvmsg_locked						
+ 1,11% :18102	[kernel.kallsyms] [k] tcp_rcv_established	+					[k] skb_try_coalesce						
+ 1,09% swapper	[kernel.kallsyms] [k] skb_gro_receive	+					[k] tcp_v4_rcv						
+ 0,87% swapper	[kernel.kallsyms] [k]inet_lookup_established	+					[k] tcp_v4_do_rcv						
+ 0,85% :18102	[kernel.kallsyms] [k] tcp_rcv_space_adjust	+					[k]inet_lookup_established						
+ 0,84% swapper	[kernel.kallsyms] [k]napi_build_skb	+					[k] check_preemption_disabled						
+ 0,82% swapper	[kernel.kallsyms] [k] intel_idle	+					[k] sock_rfree						
+ 0,82% :18102	[kernel.kallsyms] [k] skb_try_coalesce		0,69%				[k] net_rx_action						
uip: When collecting LBR b	cktraces usestitch-lbr to handle more than 32 deep entries: perf recordcall-graph lbr ; perf reportstitc	[]ip:	: Compar	re pert	formance results wi	ith: perf diff	[ <old file=""> <new file="">]</new></old>						

田     Devmem - 1 CPU (74.39 Gbps)	田 Devmem - 1 CPU - HW GRO (131.25 Gbps)
Samples: 100K of event 'cpu-cycles', Event count (approx.): 93207866858	Samples: 100K of event 'cpu-cycles', Event count (approx.): 93160176214
Overhead Command Shared Object Symbol	Overhead Command Shared Object Symbol
+ 8,02% :8076 [kernel.vmlinux] [k] gen_pool_has_addr	+ 9,33% :8892 [kernel.vmlinux] [k] tcp_recvmsg_dmabuf
+ 7,88% :8076 [kernel.vmlinux] [k] gen_pool_free_owner	+ 7,52% :8892 [kernel.vmlinux] [k] put_cmsg
+ 7,70% :8076 [idpf] [k] idpf_vport_splitq_napi_poll	+ 6,57% :8892 [kernel.vmlinux] [k] xas_store
+ 7,04% :8076 [kernel.vmlinux] [k] tcp_recvmsg_dmabuf	+ 4,74% :8892 [idpf] [k] idpf_vport_splitq_napi_poll
+ 6,52% :8076 [kernel.vmlinux] [k] tcp_gro_receive	+ 4,49% :8892 [kernel.vmlinux] [k] xas_find_marked
+ 6,48% :8076 [kernel.vmlinux] [k] gen_pool_alloc_algo_owner	+ 3,97% :8892 [kernel.vmlinux] [k] xas_load
+ 4,25% :8076 [kernel.vmlinux] [k] put_cmsg	+ 3,81% :8892 [kernel.vmlinux] [k] napi_pp_put_page
+ 3,65% :8076 [kernel.vmlinux] [k] xas_store	+ 3,30% :8892 [libeth] [k] libeth_rx_recycle_slow
+ 3,27% :8076 [kernel.vmlinux] [k] napi_pp_put_page	+ 2,95% :8892 [kernel.vmlinux] [k] _raw_spin_lock_bh
+ 2,34% :8076 [kernel.vmlinux] [k] xas_find_marked	+ 2,27% :8892 [kernel.vmlinux] [k] page_pool_alloc_frag_netmem
+ 2,28% :8076 [kernel.vmlinux] [k] xas_load	+ 2,26% :8892 [kernel.vmlinux] [k] xas_clear_mark
+ 1,75% :8076 [kernel.vmlinux] [k] _raw_spin_lock_bh	+ 1,74% :8892 [kernel.vmlinux] [k] check_preemption_disabled
+ 1,65% :8076 [kernel.vmlinux] [k] check_preemption_disabled	+ 1,42% :8892 [kernel.vmlinux] [k] net_is_devmem_iov
+ 1,56% :8076 [kernel.vmlinux] [k] dev_gro_receive	+ 1,42% :8892 [kernel.vmlinux] [k]xa_cmpxchg
+ 1,30% :8076 [kernel.vmlinux] [k] xas_clear_mark	+ 1,26% :8892 [kernel.vmlinux] [k] xas_start
+ 1,26% :8076 [kernel.vmlinux] [k] net_is_devmem_iov	+ 1,24% :8892 [kernel.vmlinux] [k] page_pool_put_unrefed_netmem
+ 1,20% :8076 [kernel.vmlinux] [k] page_pool_alloc_frag_netmem	+ 1,05% :8892 [kernel.vmlinux] [k] preempt_count_add
+ 1,09% :8076 [kernel.vmlinux] [k] inet_gro_receive	+ 1,01% :8892 [kernel.vmlinux] [k] tcp_gro_receive
+ 1,00% :8076 [kernel.vmlinux] [k] skb_gro_receive	+ 0,96% :8892 [kernel.vmlinux] [k] xas_create
+ 0,89% :8076 [kernel.vmlinux] [k]napi_build_skb	+ 0,94% :8892 [kernel.vmlinux] [k]xa_alloc
+ 0,88% :8076 [kernel.vmlinux] [k] eth_type_trans	+ 0,89% :8892 [kernel.vmlinux] [k] xas_set_mark
+ 0,85% :8076 [kernel.vmlinux] [k] page_pool_put_unrefed_netmem	+ 0,86% :8892 [kernel.vmlinux] [k]local_bh_enable_ip
+ 0,74% :8076 [kernel.vmlinux] [k] napi_build_skb	+ 0,82% :8892 [kernel.vmlinux] [k] page_pool_refill_alloc_cache

### Syscall Costs (All In One Core)

#### :10435 (10435), 1627807 events, 100.0%

syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
			·				
setsockopt	313019	0	63395.250	0.001	0.203	460.799	0.84%
recvmsg	168459	1391	26726.717	0.001	0.159	84.382	0.32%
ioctl	334018	0	1547.941	0.001	0.005	26.047	3.12%
epoll_wait	1398	0	727.045	0.000	0.520	100.199	25.99%

tcpzc2 (15989), 2172455 events, 100.0%

syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
io_uring_enter	1087324	0	46435.945	0.001	0.043	42.240	1.09%

### Devmem Syscall Costs (2 vs 1 CPU)

:14342 (14342), 4745666 events, 100.0%

syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
recvmsg	615359	1830 6	68831.493	0.001	0.112	0.371	0.03%
setsockopt	530116	0 2	22841.143	0.001	0.043	0.284	0.03%
ioctl	1224901	0	1232.993	0.001	0.001	0.604	0.09%
epoll_wait	1835	0	628.492	0.000	0.343	100.162	28.36%

2 CPUs

1 CPU

:10435 (10435), 1627807 events, 100.0%

syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
			·				
setsockopt	313019	0	63395.250	0.001	0.203	460.799	0.84%
recvmsg	168459	1391	26726.717	0.001	0.159	84.382	0.32%
ioctl	334018	0	1547.941	0.001	0.005	26.047	3.12%
epoll_wait	1398	0	727.045	0.000	0.520	100.199	25.99%

### Devmem Recycling Cost

田	Devmem - 2 CPUs (118.47 Gbps)	民	Devmem - 2 CPUs - HW GRO (111.95 Gbps)	
Samples: 199K of event	'cpu-cycles', Event count (approx.): 168906025843	Samples: 198K of ev	ent 'cpu-cycles', Event count (approx.): 142019659159	
Overhead Command	Shared Object Symbol	Overhead Command	Shared Object Symbol	
+ 18,61% :18102	[kernel.kallsyms] [k] tcp_recvmsg_dmabuf	+ 22,57% :18805	[kernel.kallsyms] [k] tcp_recvmsg_dmabuf	
+ 0,54% swapper	[idpf] [k] idpf_vport_splitg_napi_poll	+ 8,54% :18805	[kernel.kallsyms] [k] napi_pp_put_page	
+ 6,29% :18102	[kernel.kallsyms] [k] napi_pp_put_page	+ 7,05% swapper	[idpf] [k] idpf_vport_splitq_napi_poll	
+ 5,71% swapper	[kernel.kallsyms] [k] tcp_gro_receive	+ 4,31% swapper	[libeth] [k] libeth_rx_recycle_slow	
+ 5,21% swapper	[kernel.kallsyms] [k] napi_pp_put_page	+ 4,12% :18805	[kernel.kallsyms] [k] put_cmsg	
+ 3,63% :18102	[kernel.kallsyms] [k] put_cmsg	+ 3,87% swapper	[kernel.kallsyms] [k] napi_pp_put_page	
+ 2,93% :18102	[kernel.kallsyms] [k] xas_store	+ 3,41% :18805	[kernel.kallsyms] [k] xas_store	
+ 2,12% :18102	[kernel.kallsyms] [k] xas_find_marked	+ 2,64% swapper	[kernel.kallsyms] [k] intel_idle	
+ 1,82% :18102	[kernel.kallsyms] [k] xas_load	+ 2,37% :18805	[kernel.kallsyms] [k] xas_find_marked	
+ 1,68% :18102	[kernel.kallsyms] [k] _raw_spin_lock_bh	+ 2,05% :18805	[kernel.kallsyms] [k] xas_load	
+ 1,53% swapper	[kernel.kallsyms] [k] napi_build_skb	+ 1,87% :18805	[kernel.kallsyms] [k]_raw_spin_lock_bh	
+ 1,32% swapper	[kernel.kallsyms] [k] tcp_v4_rcv	+ 1,46% :18805	[kernel.kallsyms] [k] tcp_rcv_established	
+ 1,31% swapper	[kernel.kallsyms] [k] dev_gro_receive	+ 1,37% swapper	[kernel.kallsyms] [k] napi_build_skb	
+ 1,26% swapper + 1,24% :18102	[kernel.kallsyms] [k] skb_release_data [kernel.kallsyms] [k] tcp_recvmsg_locked	+ 1,25% swapper + 1,25% :18805	[kernel.kallsyms] [k] skb_release_data	
+ 1,12% :18102	[kernel.kaltsyms] [k] xas_lear_mark	+ 1,23% :18805	[kernel.kallsyms] [k] xas_clear_mark [kernel.kallsyms] [k] tcp_recvmsg_locked	
+ 1,11% :18102	[kernel.kalisyms] [k] tcp_rcv_established	+ 1,12% :18805	[kernel.kallsyms] [k] skb_try_coalesce	
+ 1,09% swapper	[kernel.kallsyms] [k] skb_gro_receive	+ 1,07% swapper	[kernel.kallsyms] [k] sko_t/y_Coateste	
+ 0,87% swapper	[kernel.kallsyms] [k]_incellookup_established	+ 1,01% :18805	[kernel.kallsyms] [k] tcp_v4_do_rcv	
+ 0,85% :18102	[kernel.kallsyms] [k] tcp_rcv_space_adjust	+ 0,79% swapper	[kernel.kallsyms] [k]inet_lookup_established	
+ 0,84% swapper	[kernel.kallsyms] [k]napi_build.skb	+ 0,72% :18805	[kernel.kallsyms] [k] check_preemption_disabled	
+ 0,82% swapper	[kernel.kallsyms] [k] intel.idle	+ 0,69% :18805	[kernel.kallsyms] [k] sock_rfree	
+ 0,82% :18102	[kernel.kallsyms] [k] skb_try_coalesce	+ 0,69% swapper	[kernel.kallsyms] [k] net_rx_action	
	R backtraces usestitch-lbr to handle more than 32 deep entries: perf recordcall-graph lbr ; perf reportst	itc Tin: Compare perfor	mance results with: perf diff [cold file> <new file="">]</new>	
		Ith compare perror		
₽.	Devmem - 1 CPU (74.39 Gbps)		Devmem - 1 CPU - HW GRO (131.25 Gbps)	
₽ Samples: 100K of event	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93287866858	Samples: 100K of ev	Devrmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214	
B Samples: 100K of event Overhead Command	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol	Samples: 100K of ev Overhead Command	Devmem - 1 CPU - HW GRO(131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol	
B Samples: 100K of event Overhead Command + 8,02% :8076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 92207866858 Shared Object Symbol [kernel.wwllnux] [ky] gen_pool_has_addr	Samples: 100K of ev Overhead Command + 9,355 :8892	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93166176214 Shared Object Symbol [kernel.wmlinux] [k] [cp.recvmsg_dmabuf	
B         Image: Samples: 100K of event           Overhead         Command           +         8,02%         :8076           +         7,88%         :8076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol [kernel.vmlinux] [k] gen_pool_has_addr [kernel.vmlinux] [k] gen_pool_free_owner	Samples: 100K of ev Overhead Command + 9,355 :8892 + 7,521 :8892	Devmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.vmlinux] [k] ptc_cncymsg_dmabuf [kernel.vmlinux] [k] ptc_cnsg	
P           Samples: 100K of event           Overhead         Command           +         0,025           :8076           +         7,80%           :8076           +         7,7%           :8076	Devmem-1CPU(74.39 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol [kernel.vmlinux] [kj gen_pool_has_addr [kernel.vmlinux] [kj dpf_uport.splitq_napi_poll [idpf] [k] idpf_uport.splitq_napi_poll	Samples: 100K of ev overhead Command + 9,355 :8892 + 7,524 :8892 + 6,574 :8892	Dewmerm - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.wmlinux] [k] tcp.recvnsg.dmabuf [kernel.wmlinux] [k] put_cmsg [kernel.wmlinux] [k] xas_store	
B Bamples: 100K of event Overhead command + 2,025 :8076 + 7,205 :8076 + 7,205 :8076 + 7,045 :8076	Devmem - 1 CPU (74.39 Gbps) 'pu-cycles', Event count (approx.): 93287866858 Shared Object Symbol [kernel.vmlitux] [k] gen_pool.has.addr [kernel.vmlitux] [k] gen_pool.free_owner [tdpf] [k] tdpf_vport_splitg.napi.poll [kernel.vmlitux] [k] tcp_recvsg_dmabuf	Samples: 100K of ev Overhead Command + 0,35% :8892 + 7,52% :8892 + 6,57% :8892 + 4,74% :8892	Dewmerm - 1 CPU - HW GRC0(131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93569176214 Shared Object Symbol [kernel.wmlinux] [k] tcp_recvnsg_dmabuf [kernel.wmlinux] [k] xas_store [kernel.wmlinux] [k] xas_store [idpf] [k] idpfyport_splitg_napi_poll	
B           Samples: 100K of event           Overhead         Command           +         0,02%           >:8076           +         7,05%           :8076           +         7,05%           :8076           +         7,04%           :8076           +         0,52%           :8076	Devmem-1CPU(7439 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 shared Object Symbol [kernel.vmlinux] [ki gen_pool_has_addr [kernel.vmlinux] [ki dpf_uport.splitq_napt_poll [kernel.vmlinux] [ki tcp_recvssg_dmabuf [kernel.vmlinux] [ki tcp_recevel	Samples: 100K of ev Overhead Command + 0,35% :8092 + 7,52% :8092 + 6,57% :8092 + 4,74% :8092 + 4,45% :8092	Dewmerm - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.vmllinux] [k] toc.precvnsg_dmabuf [kernel.vmllinux] [k] put_cmsg [kernel.vmllinux] [k] xas_store [idpf] [k] idpf_yport_splitg_napi_poll [kernel.vmllinux] [k] xas_iton_marked	
B         Samples: 100K of event           Overhead         Command           0.2         :8076           7,805         :8076           7,705         :8076           6,522         :8076           6,524         :8076           6,524         :8076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93287866858 Shared Object Symbol [kernel.vmlitux] [k] gen pool.has.addr [kernel.vmlitux] [k] gen pool.free owner [tdpf] [k] top.recvsg.genabut [kernel.vmlitux] [k] top.recvsg.genabut [kernel.vmlitux] [k] top.grc.recetve [kernel.vmlitux] [k] gen_pool.alloc_algo_owner	Samples: 100K of ev Dverhead Command + 0,155: 18092 + 7,552: 18092 + 7,557: 18092 + 4,74%: 18092 + 4,49%: 18092 + 3,97%: 18092	Dewmerm - 1 CPU - HW GRC0(131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.wmlinux] [k] tcp_recvnsg_dnabuf [kernel.wmlinux] [k] xas_store [idpf] [k] idf/yport_splitq_napi_poll [kernel.wmlinux] [k] xas_find_marked [kernel.wmlinux] [k] xas_load	
H         Samples: 108K of event           Overhead         Command           0.22         :8076           7.08%         :8076           7.70%         :8076           7.22         :8076           6.22         :8076           6.22         :8076           6.45%         :8076           4.25%         :8076	Devmem-1CPU(7439 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared 00ject Symbol [kernel.wnlinux] [ki gen_pool_has_addr [kernel.wnlinux] [ki dpf_uport.splitq_napt_poll [kernel.wnlinux] [ki tcp_precvsg_dmabuf [kernel.wnlinux] [ki tcp_precvedue [kernel.wnlinux] [ki tcp_or_cectve [kernel.wnlinux] [ki gen_pool_alloc_algo_owner [kernel.wnlinux] [ki put_msg	Samples: 100K of ev overhead Command + 7,55%: 8892 + 7,55%: 8892 + 6,57%: 8892 + 4,74%: 8892 + 4,74%: 8892 + 3,97%: 8892 + 3,95%: 8892	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 [kernel.vmllnux] [k] tcp.recvnsg_dmabuf [kernel.vmllnux] [k] put_cmsg [kernel.vmllnux] [k] xas_store [idpf] [k] idpf_yport_splitq_napi_poll [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load	
H         Samples: 100K of event           Overhead Command         0076           + 7,00%         18076           + 7,70%         18076           + 7,00%         18076           + 7,00%         18076           + 6,52%         18076           - 6,52%         18076           - 4,25%         18076           - 3,05%         18076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93287866858 Shared Object Symbol [kernel.wnlinux] [k] gen pool.has.addr [kernel.wnlinux] [k] gen pool.free owner [tdpf] [k] top.revssg.gdmabuf [kernel.wnlinux] [k] top.gre.recetve [kernel.wnlinux] [k] gen_pool.alloc_algo_owner [kernel.wnlinux] [k] gen_pool.alloc_algo_owner [kernel.wnlinux] [k] xas.tore	Samples: 100K of ev.           Samples: 100K of ev.           0yerhead Command           + 7,553           + 7,553           + 6,577           + 6,777           + 6,777           + 4,49%           + 8892           + 3,757           + 4,49%           + 8892           + 3,95%           + 3,95%           + 3,95%           + 3,95%           + 3,95%	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 9366176214 Shared Object Symbol [kernel.wmlinux] [k] tcp_recvnsg_dnabuf [kernel.wmlinux] [k] xas_store [idpf] [k] idpf.yoprt_splitq_napi_poll [kernel.wmlinux] [k] xas_find_marked [kernel.wmlinux] [k] xas_load [kernel.wmlinux] [k] xas_load [kernel.wmlinux] [k] xas_load	
BE         Samples: 100K of event           Overhead Command	Devmem-1CPU(7439 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol [kernel.vmlinux] [ki gen_pool_has_addr [kernel.vmlinux] [ki dpf_uport.splitq_napi_poll [kernel.vmlinux] [ki tdp_recvsg_dmabuf [kernel.vmlinux] [ki tcp_pre_receive [kernel.vmlinux] [ki tcp_ore_edive [kernel.vmlinux] [kj put_engg [kernel.vmlinux] [kj put_engg [kernel.vmlinux] [kj napi_popl_tage	Samples:         100K of ev Overhead         Command           * 7,55%         :0852         *	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.vmlinux] [k] tcc.precvnsg_dmabuf [kernel.vmlinux] [k] xas_store [kipf] [k] tidpf_yport_splitq_napi_poll [kernel.vmlinux] [k] xas_itod_marked [kernel.vmlinux] [k] xas_load [kernel.vmlinux] [k] xas_load [kernel.vmlinux] [k] tbch_rx_recycle_slow [kibeth] [k] [k] cs.bh	
H         Samples:         100K of event           Overhead Command         6076           *         7,070         18076           *         7,070         18076           *         7,070         18076           *         7,087         18076           *         6,452         18076           *         6,555         18076           *         3,255         18076           *         3,243         18076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93287866858 Shared Object Symbol [kernel.wnlinux] [k] gen pool.has.addr [kernel.wnlinux] [k] gen pool.free owner [tdpf] [k] top.revsesg.genabut [kernel.wnlinux] [k] top.grev.resetve [kernel.wnlinux] [k] gen.pool.alloc.algo_owner [kernel.wnlinux] [k] gen.pool.alloc.algo_owner [kernel.wnlinux] [k] xas.tore [kernel.wnlinux] [k] xas.tore [kernel.wnlinux] [k] xas.tore [kernel.wnlinux] [k] xas.tore	Samples:         10% of ex           Overhead         Command           y 15%         18892           + 7,55%         18892           + 6,57%         18892           + 4,45%         18892           + 3,81%         18892           + 3,81%         18892           + 3,95%         18892           + 3,95%         18892           + 3,95%         18892           + 2,95%         18892           + 2,77%         18892	Dewmerm - 1 CPU - HW GRO(131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 8360176214 Shared Object Symbol [kernel.wmlinux] [k] tcp_recvmsg_dnabuf [kernel.wmlinux] [k] xas_store [idpf] [k] idpf.yoprt_splitq_napi_poll [kernel.wmlinux] [k] xas_find_marked [kernel.wmlinux] [k] xas_load [kernel.wmlinux] [k] xas_load [kernel.wmlinux] [k] app.put_page [libth] [k] raw_spln_lock_bh [kernel.wmlinux] [k] _raw_spln_lock_bh	
BE         Samples: 100K of event           Overhead Command	Devmem-1CPU(7439 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol [kernel.vmlinux] [ki gen_pool_has_addr [kernel.vmlinux] [ki dpf_uport.splitq_napi_poll [kernel.vmlinux] [ki tdp_recvsg_dmabuf [kernel.vmlinux] [ki tcp_pre_receive [kernel.vmlinux] [ki tcp_ore_edive [kernel.vmlinux] [kj put_engg [kernel.vmlinux] [kj put_engg [kernel.vmlinux] [kj napi_popl_tage	Samples:         100K of ev Overhead         Command           * 7,55%         :0852         *	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.vmlinux] [k] tcc.precvnsg_dmabuf [kernel.vmlinux] [k] xas_store [idpf] [k] tdpf_yport_splitq_napi_poll [kernel.vmlinux] [k] xas_itod_marked [kernel.vmlinux] [k] xas_load [kernel.vmlinux] [k] xas_load [kernel.vmlinux] [k] tbch_rx_recycle_slow [kibeth] [k] [k] cs.bh	
BE         Samples: 100K of event           Overhead Command	Devmem-1CPU(7439 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol [kernel.wmiinux] [k] gen_pool_has_addr [kernel.wmiinux] [k] gen_pool_has_addr [kernel.wmiinux] [k] tdpr_port.splitq_napi_poll [kernel.wmiinux] [k] tcp_pr_cresys_dmabuf [kernel.wmiinux] [k] tcp_pr_crecelve [kernel.wmiinux] [k] put_cnsg [kernel.wmiinux] [k] put_sg [kernel.wmiinux] [k] napi_pp_put_page [kernel.wmiinux] [k] napi_pp_put_page [kernel.wmiinux] [k] xas_find_marked [kernel.wmiinux] [k] xas_find_marked	Constraint         Constraint           Samples:         100K of ev overhead         Command           +         7,55k         10852           +         7,55k         10852           +         4,74k         10852           +         3,75k         10852           +         4,474k         10852           +         3,10k         10892           +         3,10k         10892           +         2,575k         10892           +         2,75k         10892           +         2,275k         10892           +         2,275k         10892           +         2,27k         10892	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.vmllnux] [k] tcp_recvnsg_dmabuf [kernel.vmllnux] [k] put_cnsg [kernel.vmllnux] [k] xas_store [idpf] [k] tdpf_yport_splitq_napi_poll [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] nas_p.put_page [libeth] [k] libeth_rx_recycle_slow [kernel.vmllnux] [k] page_pool_alloc_frag_netmem [kernel.vmllnux] [k] yas_load [kernel.vmllnux] [k] page_pool_alloc_frag_netmem	
H         Samples:         100K of event           Supples:         100K of event           Overhead Command         50076           7,000:         18076           7,700:         18076           7,000:         18076           6,522:         18076           4,255:         18076           3,055:         18076           2,344:         18076           2,255:         18076           2,255:         18076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93287866858 Shared Object Symbol [kernel.wnlinux] [k] gen.pool.has.addr [kernel.wnlinux] [k] cp.pocv.sg.gdmabuf [kernel.wnlinux] [k] tcp.prcvsg.gdmabuf [kernel.wnlinux] [k] tcp.prcvsg.gdmabuf [kernel.wnlinux] [k] tcp.prcvsg.gdmabuf [kernel.wnlinux] [k] gen.pool.alloc.algo_aowner [kernel.wnlinux] [k] xas.store [kernel.wnlinux] [k] xas.store	Samples:         10% of ex           Overhead         Commands           0,25%         18892           +         7,5%           +         7,5%           +         4,4%           8892         +           +         7,5%           +         4,4%           8892         +           +         3,81%           *         3,9%           *         3,9%           *         2,9%           *         2,9%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         2,2%           *         3,1%           *         3,2%           *         3,2%           *         3,2%           *         3,2%           *         3,2%           *         3,2%           *         3,2%           *	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 8360176214 Shared Object Symbol [kernel.wmlinux] [k] tcp.recvnsg.dnabuf [kernel.wmlinux] [k] txp.recvnsg.dnabuf [kernel.wmlinux] [k] xas.store [idpf] [k] idpf.yoprt.splitq.napi_poll [kernel.wmlinux] [k] xas.find.marked [kernel.wmlinux] [k] xas.joad [kernel.wmlinux] [k] xas.joad [kernel.wmlinux] [k] aras.gol.ad [kernel.wmlinux] [k] aras.gol.adloc.frag.netmem [kernel.wmlinux] [k] xas.ctear.mark [kernel.wmlinux] [k] xas.ctear.mark [kernel.wmlinux] [k] zas.ctear.mark	
BE         Samples: 100K of event           Overhead Command         Command           + 0.02::8076         :8076           + 7,00%         :8076           + 7,04%         :8076           + 7,04%         :8076           + 6,322::8076         :8076           - 3,25%         :8076           - 3,25%         :8076           - 3,25%         :8076           - 1,75%         :8076           + 1,75%         :8076	Devmem-1CPU(7439 Gbps) 'cpu-cycles', Event count (approx.): 93207866858 Shared Object Symbol [kernel.wmiinux] [ki gen_pool_has_addr [kernel.wmiinux] [ki gen_pool_has_addr [tdpf] [kernel.wmiinux] [ki tdpr_urct.splitq_napi_poll [kernel.wmiinux] [ki tcp_or_cecive [kernel.wmiinux] [ki gen_pool_alloc_algo_owner [kernel.wmiinux] [ki put_cmsg [kernel.wmiinux] [ki napi_pp_ut_page [kernel.wmiinux] [ki napi_pp_ut_page [kernel.wmiinux] [ki xas_find_marked [kernel.wmiinux] [ki xas_find_marked [kernel.wmiinux] [ki xas_find_marked [kernel.wmiinux] [ki xas_find_d [kernel.wmiinux] [ki xas_find_d [kernel.wmiinux] [ki check_premption_disabled	Samples:         100K of ev Overhead         Command           + 7,55k         :0852         +         -	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93160176214 Shared Object Symbol [kernel.vmllnux] [k] toc.recvnsg.dmabuf [kernel.vmllnux] [k] put_cnsg [kernel.vmllnux] [k] xas.store [idpf] [k] tidpf_yport_splitq_napi_poll [kernel.vmllnux] [k] xas.load [kernel.vmllnux] [k] xas.load [kernel.vmllnux] [k] xas.load [kernel.vmllnux] [k] pap.put.page [libeth] [k] libeth_rx_rcycle_slow [kernel.vmllnux] [k] page.pool_alloc_frag.netmem [kernel.vmllnux] [k] xas.load [kernel.vmllnux] [k] page.pool_alloc_frag.netmem [kernel.vmllnux] [k] xas.idevnem, iov	
H         Samples:         100K of event           Overhead Command         6076         6076           - 7,000:         18076         7,000:         18076           - 7,000:         18076         6,052         18076           - 7,000:         18076         6,0576         3,055         18076           - 3,255:         18076         3,275         18076           - 2,343:         18076         2,285         18076           - 1,055:         18076         1,055         18076           - 1,055:         18076         1,055         13076           - 1,055:         18076         1,055         13076           - 1,055:         18076         1,055         13076	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 93287866858 Shared Object Symbol [kernel.wnlinux] [k] gen_pool.has.addr [kernel.wnlinux] [k] cpo.cpc.spltg.napt.poll [kernel.wnlinux] [k] tcp_rcevsg.gdmabuf [kernel.wnlinux] [k] tcp_rcevsg.gdmabuf [kernel.wnlinux] [k] tcp_rcevsg.gdmabuf [kernel.wnlinux] [k] gen_pool.alloc.algo_aowner [kernel.wnlinux] [k] xas.store [kernel.wnlinux] [k] (k] (k] yergengtion_disabled [kernel.wnlinux] [k] (dev_gor_creetive	Samples:         10% of ex           Overhead         Commands           0,3%         18892           +         0,5%           +         0,5%           +         0,5%           +         0,5%           +         0,5%           +         0,5%           +         0,4%           +         0,5%           +         3,6%           +         3,36%           +         2,9%           +         2,7%           +         2,2%           +         2,2%           +         2,2%           +         1,4%           +         1,4%           +         1,4%           +         1,4%           +         1,4%           +         1,4%           +         1,4%	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 8360176214 Shared Object Symbol [kernel.wmlinux] [k] tcp.recvnsg.dnabuf [kernel.wmlinux] [k] xas.store [idpf] [k] idpf.yoprt.splitq.napi_poll [kernel.wmlinux] [k] xas.ftnd.marked [kernel.wmlinux] [k] xas.flod.marked [kernel.wmlinux] [k] xas.gload [kernel.wmlinux] [k] xas.pload [kernel.wmlinux] [k] xas.pload [kernel.wmlinux] [k] xas.pload [kernel.wmlinux] [k] xas.pload. [kernel.wmlinux] [k] zas.pol.alloc.frag.netmem [kernel.wmlinux] [k] xas.clear.mark [kernel.wmlinux] [k] xas.clear.mark [kernel.wmlinux] [k] zas.clear.mark [kernel.wmlinux] [k] zas.clear.mark [kernel.wmlinux] [k] zas.cmerem.jov [kernel.wmlinux] [k] zas.cmerem.jov	
BE         Samples: 100K of event           Overhead Command         Command           + 0.02x: 18076         18076           + 7,00%: 18076         18076           + 7,00%: 18076         18076           + 7,00%: 18076         18076           - 7,00%: 18076         3,25%: 18076           - 3,25%: 18076         2,34%: 18076           - 2,34%: 18076         1,75%: 18076           - 1,55%: 18076         1,55%: 18076	Devmem-1CPU(7439 Gbps)         'cpu-cycles', Event count (approx.): 93207066858         Shared Object         Symbol         [kernel.wmitnux]         [k] gen_pool_has_addr         [kernel.wmitnux]         [k] gen_pool_free_owner         [tdf]         [kernel.wmitnux]         [k] dpf_upcrt.splitq_napi_poll         [kernel.wmitnux]         [k] tcp_precvssg_dmabuf         [kernel.wmitnux]         [k] tcp_precvsg_dmabuf         [kernel.wmitnux]         [k] tcp_precvsg_dmabuf         [kernel.wmitnux]         [k] tcp_precvesg_dmabuf         [kernel.wmitnux]         [k] tcp_precvesg_dmabuf         [kernel.wmitnux]         [k] pol_lalo_algo_owner         [kernel.wmitnux]         [k] pol_lalog_algo_owner         [kernel.wmitnux]         [k] pol_tasg         [kernel.wmitnux]         [k] mas_strind_marked         [kernel.wmitnux]         [k] raw_spin_lock_bh         [kernel.wmitnux]         [k] check_preception_disabled         [kernel.wmitnux]         [k] check_preception_disabled         [kernel.wmitnux]         [k] page_pool_alloc_frag_netmen </td <td>Samples:         100K of ev Overhead         Command           +         7,55%         16852           +         7,55%         16852           +         7,55%         16852           +         4,74%         16852           +         4,74%         16852           +         3,10%         16892           +         3,10%         16892           2         2,77%         16892           2         2,75%         16892           2         2,75%         16892           2         2,75%         16892           2         2,75%         16892           2         2,75%         16892           2         1,74%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892</td> <td>Dewmem - 1 CPU - HW GRO (131.25 Gbps)           ant 'cpu-cycles', Event count (approx.): 93169176214           Symbol           [kernel.vmllunx] [k] tCp_recvmsg_dmabuf           [kernel.vmllunx] [k] put_cmsg           [kernel.vmllunx] [k] put_cmsg           [kernel.vmllunx] [k] xas_store           [tdpf]           [k] tdpf_yport_splitq_napi_poll           [kernel.vmllunx] [k] xas_load           [kernel.vmllunx] [k] xas_load           [kernel.vmllunx] [k] tbbch_rx_recycle_slow           [ltbcth]         [k] tbbch_rx_recycle_slow           [kernel.vmllunx] [k] yas_pln_lock/bn           [kernel.vmllunx] [k] xas_load           [kernel.vmllunx] [k] zs_pln_lock/bn           [kernel.vmllunx] [k] zs_pln_lock/bn           [kernel.vmllunx] [k] yage_pol_alloc_frag_netmem           [kernel.vmllunx] [k] xas_idevmem_iov           [kernel.vmllunx] [k] zs_devme_iov           [kernel.vmllunx] [k] zs_start           [kernel.vmllunx] [k] yage_pol_pol_ut_unrefed_netmem           [kernel.vmllunx] [k] yage_pol_ol_alturefed_netmem</td> <td></td>	Samples:         100K of ev Overhead         Command           +         7,55%         16852           +         7,55%         16852           +         7,55%         16852           +         4,74%         16852           +         4,74%         16852           +         3,10%         16892           +         3,10%         16892           2         2,77%         16892           2         2,75%         16892           2         2,75%         16892           2         2,75%         16892           2         2,75%         16892           2         2,75%         16892           2         1,74%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892           1         1,42%         16892	Dewmem - 1 CPU - HW GRO (131.25 Gbps)           ant 'cpu-cycles', Event count (approx.): 93169176214           Symbol           [kernel.vmllunx] [k] tCp_recvmsg_dmabuf           [kernel.vmllunx] [k] put_cmsg           [kernel.vmllunx] [k] put_cmsg           [kernel.vmllunx] [k] xas_store           [tdpf]           [k] tdpf_yport_splitq_napi_poll           [kernel.vmllunx] [k] xas_load           [kernel.vmllunx] [k] xas_load           [kernel.vmllunx] [k] tbbch_rx_recycle_slow           [ltbcth]         [k] tbbch_rx_recycle_slow           [kernel.vmllunx] [k] yas_pln_lock/bn           [kernel.vmllunx] [k] xas_load           [kernel.vmllunx] [k] zs_pln_lock/bn           [kernel.vmllunx] [k] zs_pln_lock/bn           [kernel.vmllunx] [k] yage_pol_alloc_frag_netmem           [kernel.vmllunx] [k] xas_idevmem_iov           [kernel.vmllunx] [k] zs_devme_iov           [kernel.vmllunx] [k] zs_start           [kernel.vmllunx] [k] yage_pol_pol_ut_unrefed_netmem           [kernel.vmllunx] [k] yage_pol_ol_alturefed_netmem	
H         Samples:         100K of event           Overhead Command         0.007.00         0.007.00           Y, 200:         1007.0         1007.0           Y, 201:         1007.0         1007.0	Devmem - 1 CPU (74.39 Gbps) 'cpu-cycles', Event count (approx.): 9320766050 (kernel.wnlinux) [k] gen_pool_has_addr [kernel.wnlinux] [k] gen_pool_free_owner [tdpf] [k] tcp_rcevsg_gdmabuf [kernel.wnlinux] [k] tcp_rcevsg_gdmabuf [kernel.wnlinux] [k] tcp_rcevsg_gdmabuf [kernel.wnlinux] [k] tcp_rcevsg_gdmabuf [kernel.wnlinux] [k] gen_pool_alloc_algo_avmer [kernel.wnlinux] [k] xas_ftord [kernel.wnlinux] [k] check_preemption_disabled [kernel.wnlinux] [k] kas_clear_mark [kernel.wnlinux] [k] net_sig_even_fov [kernel.wnlinux] [k] net_sig_even_fov [kernel.wnlinux] [k] net_sig_even_fov [kernel.wnlinux] [k] net_sig_even_fov	Samples:         10%         of exponent of expon	Dewmern - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 3360176214 Shared Object Symbol [kernel.wmlinux] [k] tcp_recvmsg_dnabut [kernel.wmlinux] [k] xas_store [idpf] [k] idpf/yoprt_splitq_napi_poll [kernel.wmlinux] [k] xas_ftnd_marked [kernel.wmlinux] [k] xas_load [kernel.wmlinux] [k] xas_load [kernel.wmlinux] [k] xas_pol_alloc_frag_netmem [kernel.wmlinux] [k] xas_pol_alloc_frag_netmem [kernel.wmlinux] [k] xas_ctear.mark [kernel.wmlinux] [k] xas_ctear.mark [kernel.wmlinux] [k] xas_ctear.mark [kernel.wmlinux] [k] net_is_dewmem_iov [kernel.wmlinux] [k] net_is_dewmem_iov [kernel.wmlinux] [k] xas_ctart [kernel.wmlinux] [k] xas_start [kernel.wmlinux] [k] prempt_count_add [kernel.wmlinux] [k] prempt_count_add [kernel.wmlinux] [k] prempt_count_add [kernel.wmlinux] [k] prempt_count_add	
B         Samples: 100K of event           Overhead Command         Command           + 0.0xx: 8076         8076           + 7,00x: 8076         8076           + 7,00x: 8076         8076           + 7,00x: 8076         8076           + 7,00x: 8076         8076           - 4,250: 8076         3,05%           - 3,05%: 8076         3,27%           - 3,05%: 8076         1,75%           - 1,05%: 8076         1,26%           - 1,05%: 8076         1,20%           - 1,05%: 8076         1,20%	Devmem-1CPU(7439 Gbps)         'cpu-cycles', Event count (approx.): 93207066858         Shared Object       Symbol         [kernel.wmitnux]       [kj gen_pool_has_addr         [kernel.wmitnux]       [kj gen_pool_free_owner         [tdf]       [ki dpf_uport.split(_napi_poll         [kernel.wmitnux]       [kj gen_pool_alloc_najt_poll         [kernel.wmitnux]       [kj tcp_procvest_gdmbuf]         [kernel.wmitnux]       [kj tcp_procective]         [kernel.wmitnux]       [kj pp_out_page]         [kernel.wmitnux]       [k] hapi_pout_page         [kernel.wmitnux]       [k] xas_tord         [kernel.wmitnux]       [k] xas_trind_marked         [kernel.wmitnux]       [k] chck_preception_disabled         [kernel.wmitnux]       [k] chck_preception_disabled         [kernel.wmitnux]       [k] chck_preception_disabled         [kernel.wmitnux]       [k] rea_in_ark         [kernel.wmitnux]       [k] rea_in_ark         [kernel.wmitnux]       [k] chck_preception_disabled         [kernel.wmitnux]       [k] page_pool_alloc_frag.netmem         [kernel.wmitnux]       [k] page_pool_alloc_frag.netmem         [kernel.wmitnux]       [k] sbg_or_crecetive         [kernel.wmitnux]       [k] sbg_or_crecetive         [kernel.wmitnux]	Sampler:         1000 of ev Overhead         Cormand           +         2,355         16862           +         2,557         16892           +         2,557         16892           +         2,557         16892           +         4,474         16852           +         4,474         16852           +         3,134         16852           +         3,147         16852           +         3,147         16852           +         2,279         16892           +         2,279         16892           +         2,427         16892           +         1,427         16892           +         1,247         16892           +         1,247         16892           +         1,057         16892           +         1,057         16892           +         0,967         1892	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93169176214 Swabol [kernel.vmllnux] [k] tcp_recvnsg.dmabuf [kernel.vmllnux] [k] tcp_store [kernel.vmllnux] [k] xas_store [tdpf] [k] tdpf_yport_splitq_napi_poll [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] yasp_nolckloc. [kernel.vmllnux] [k] yasp_nolckloc. [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_sin.lock/sn [kernel.vmllnux] [k] tcp_gro_recetve [kernel.vmllnux] [k] xas_sin.lock/sn [kernel.vmllnux] [k] xaspi.count_add [kernel.vmllnux] [k] xaspi.count_add	
H         Samples:         100K of event           Overhead Command         6,022         10076           7,700         10076         7,070           7,700         10076         7,070           7,700         10076         7,070           7,700         10076         7,070           7,700         10076         7,070           7,040         10076         7,075           7,055         10076         3,075           3,275         10076         2,341           2,235         10076         2,215           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055	Devmem - 1 CPU (74.39 Gbps)           'cpu-cycles', Event count (approx.): 9320706050           Shared Object Symbol           [kernel.wnlinux]         [k] gen_pool_has_addr           [kernel.wnlinux]         [k] gen_pool_free_owner           [tdp7]         [k] gen_pool_free_owner           [tdp7]         [k] tcp_rcerseg_dmabul           [kernel.wnlinux]         [k] tcs_store           [kernel.wnlinux]         [k] tcs_crecetve           [kernel.wnlinux]         [k]	Samples:         10% of exponent           Sworthead         Command           0, 35%         18892           +         0, 55%           +         0, 57%           18892         +           +         0, 57%           +         3, 61%           +         3, 36%           +         3, 36%           +         3, 36%           +         2, 57%           +         3, 36%           +         2, 27%           +         2, 27%           +         1, 36%           +         1, 42%           +         1, 42%           +         1, 26%           +         1, 26%           +         1, 26%           +         1, 65%           +         0, 96%           +         0, 96%	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 3369176214 Shared Object Symbol [kernel,wmlinux] [k] tcp.recvmsg.dmabut [kernel,wmlinux] [k] tcp.recvmsg.dmabut [kernel,wmlinux] [k] xas.store [idpf] [k] idpf.yoprt.splitg.napi_poll [kernel,wmlinux] [k] xas.flod_marked [kernel,wmlinux] [k] xas.j.oad [kernel,wmlinux] [k] xas.j.oad [kernel,wmlinux] [k] xas.j.oad [kernel,wmlinux] [k] raw.spln.lock.bh [kernel,wmlinux] [k] xas.clear.mark [kernel,wmlinux] [k] xas.clear.mark [kernel,wmlinux] [k] raw.spln.lock.fmg_netmem [kernel,wmlinux] [k] raw.spln.lock.fmg_netmem [kernel,wmlinux] [k] ras.start [kernel,wmlinux] [k] ras.start [kernel,wmlinux] [k] ras.start [kernel,wmlinux] [k] prempt_count.add [kernel,wmlinux] [k] prempt_count.add [kernel,wmlinux] [k] prempt_count.add [kernel,wmlinux] [k] prempt_count.add [kernel,wmlinux] [k] ras.stret [kernel,wmlinux] [k] ras.stret [kernel,wmlinux] [k] prempt_count.add [kernel,wmlinux] [k] ras.stret [kernel,wmlinux] [k] ras.s	
H         Samples:         106K of event           0verhead Command         Command           -         0.02:         8076           -         7.80%         8076           -         7.84%         8076           -         7.84%         8076           -         7.84%         8076           -         7.84%         8076           -         7.84%         8076           -         7.85%         8076           -         2.34%         8076           -         2.34%         8076           -         2.34%         8076           -         2.34%         8076           -         2.34%         8076           -         2.95%         8076           -         1.95%         8076           -         1.95%         8076           -         1.95%         8076           -         0.85%         8076	Devmem-1CPU(7439 Gbps)           'cpu-cycles', Event count (approx.): 93207066858           Shared Object         Symbol           [kernel.wmitnux]         [kj gen.pool_has_addr           [kernel.wmitnux]         [kj gen.pool_ree_owner           [tdf]         [ki dpf.port.splitq.napi_poll           [kernel.wmitnux]         [kj gen.pool_ree_owner           [tdf]         [kernel.wmitnux]           [kernel.wmitnux]         [kj tcp_precvsg_dmabuf           [kernel.wmitnux]         [kj tcp_precvesg_dmabuf           [kernel.wmitnux]         [kj tcp_precvesg_dmabuf           [kernel.wmitnux]         [kj tcp_precvesg_dmabuf           [kernel.wmitnux]         [kj pp_pol_alge_lalge_owner           [kernel.wmitnux]         [kj pp_pol_poge           [kernel.wmitnux]         [kj napi_pojut_page           [kernel.wmitnux]         [kj xas_tore           [kernel.wmitnux]         [kj check_preception_disabled           [kernel.wmitnux]         [k]           [kernel.wmit	Sampler:         1000 of ev Overhead         Command           +         2,355         1000 of everhead         Command           +         2,355         1000 of everhead         Command           +         2,355         1000 of everhead         Command           +         2,357         1000 of everhead         Command           +         2,357         1000 of everhead         Command           +         2,444         1000 of everhead         Command           +         1,424         1000 of everhead         Command           +         1,424         1000 of everhead         Command           +         1,445         1000 of everhead         Command           +         0,945         1000 of everhead         Command	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 93169176214 Swabol [kernel.vmllnux] [k] tcp_recvnsg_dmabuf [kernel.vmllnux] [k] xas_store [tdpf] [k] tdpf_yport_splitq_napi_poll [kernel.vmllnux] [k] xas_tore [tdpf] [k] tkg_store [tdpf] [k] theth_rx_recycle_slow [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] xas_load [kernel.vmllnux] [k] yasp_nolocklow [kernel.vmllnux] [k] yasp_nolocklow [kernel.vmllnux] [k] yasp_nolocklow [kernel.vmllnux] [k] xas_sln.lock [kernel.vmllnux] [k] xas_sln.lock [kernel.vmllnux] [k] xas_sln.lock [kernel.vmllnux] [k] xas_sln.lock [kernel.vmllnux] [k] xas_sln.lock [kernel.vmllnux] [k] xas_sln.lock [kernel.vmllnux] [k] xas_slavd [kernel.vmllnux] [k] xas_slavd [kernel.vmllnux] [k] xas_slavd [kernel.vmllnux] [k] xas_scart [kernel.vmllnux] [k] yage_pool_put_unefed_netmen [kernel.vmllnux] [k] xas_slavd [kernel.vmllnux] [k] xas_start [kernel.vmllnux] [k] xas_start	
H         Samples:         100K of event           Overhead Command         6,022         10076           7,700         10076         7,070           7,700         10076         7,070           7,700         10076         7,070           7,700         10076         7,070           7,700         10076         7,070           7,040         10076         7,075           7,055         10076         3,075           3,275         10076         2,341           2,235         10076         2,215           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055           1,055         10076         1,055	Devmem - 1 CPU (74.39 Gbps)           'cpu-cycles', Event count (approx.): 9320706050           Shared Object Symbol           [kernel.wnlinux]         [k] gen_pool_has_addr           [kernel.wnlinux]         [k] gen_pool_free_owner           [tdp7]         [k] gen_pool_free_owner           [tdp7]         [k] tcp_rcerseg_dmabul           [kernel.wnlinux]         [k] tcs_store           [kernel.wnlinux]         [k] tcs_crecetve           [kernel.wnlinux]         [k]	Samples:         10% of exponent           Sworthead         Command           0, 35%         18892           +         0, 55%           +         0, 57%           18892         +           +         0, 57%           +         3, 61%           +         3, 36%           +         3, 36%           +         3, 36%           +         2, 57%           +         3, 36%           +         2, 27%           +         2, 27%           +         1, 36%           +         1, 42%           +         1, 42%           +         1, 26%           +         1, 26%           +         1, 26%           +         1, 65%           +         0, 96%           +         0, 96%	Dewmem - 1 CPU - HW GRO (131.25 Gbps) ent 'cpu-cycles', Event count (approx.): 3369176214 Shared Object Symbol [kernel,wmlinux] [k] tcp_recvmsg_dmabut [kernel,wmlinux] [k] tcp_recvmsg_dmabut [kernel,wmlinux] [k] xas_store [idpf] [k] idpfyoprt_splitg_napi_poll [kernel,wmlinux] [k] xas_flod_marked [kernel,wmlinux] [k] xas_load [kernel,wmlinux] [k] xas_load [kernel,wmlinux] [k] xas_load [kernel,wmlinux] [k] raw.spln_lock,bh [kernel,wmlinux] [k] raw.spln_lock,chrag_netmem [kernel,wmlinux] [k] xas_clear.mark [kernel,wmlinux] [k] raw.spln_lock,frag_netmem [kernel,wmlinux] [k] raw.spln_lock,frag_netmem [kernel,wmlinux] [k] rak_sclear.mark [kernel,wmlinux] [k]	

### The IDPF HW GRO Dragon ...



HDS implies two page pools: Header and payload pool

- Simple lifecycle management: 1-1 mapping
  - IDPF uses 1 page per header page per payload
- Unfortunately when you have HW GRO, you only need 1 header for X payloads..

For sake of discussion, say we receive a GRO size of 16(\*4K payload)

- it means only the first header is relevant i.e the others are dummies
  - *libeth\_rx\_recycle\_slow()* recycles these dummy (15) header pages back to the header page pool at <u>softirq</u> context
    - foreach page {hdr producer lock, recycle, release hdr producer lock}
    - The page with header is only recycled when recvmsg() completes and skb\_release kicks in at user context (contends for hdr producer lock)\*
    - In the meantime more and more GRO'ed packets are coming in and contend for that same lock..

### Slaying The IDPF HW GRO Dragon ...

• We did not have time to slay this dragon

Potential approach:

- Keep single page for X headers and use refcounts
- For IPv4, no more than 64B is really needed
  - so 4096/64 should be much higher than max GRO size...

### The Payload Recycle Dragon



On single CPU we observed that the cost of setsockopt was higher both in the syscall trace as well as in perf

→ Recycling of the payload happens with setsockopt

After looking at the kernel code, we observed that although the uapi allows for a batch of buffers from user space to be sent to the kernel, the kernel would still do:

foreach page {grab payload producer lock, recycle, release payload producer lock}

We made two changes:

- 1. Application collects as many as X consumed buffers for recycling (max 1024)
- 2. Changed the kernel code to amortize the cost of the lock:

grab payload producer lock, for each page: recycle release payload producer lock

### Devmem RX Fix I: App buffer recycle batching

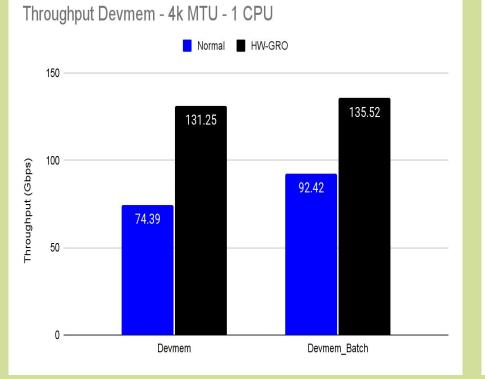
- The kernel allows up to 128 token containers with a max of up to 1024 tokens to free at once
  - The common case is only 1 token container with up to 1024 tokens
- recvmsg() will likely receive more than one token per call, up to the provided control space
  - The number of tokens will vary depending on the bandwidth and tcp buffer conditions
  - At 200G, our average token was at ~875 per recvmsg
- In our tests we accumulate small recvmsgs up to 512 tokens and then issue a setsockopt to release them
  - It reduces our setsockopt usage by 15% but did not show visible improvements on throughput a cpu utilization

# Devmem recv read size distribution

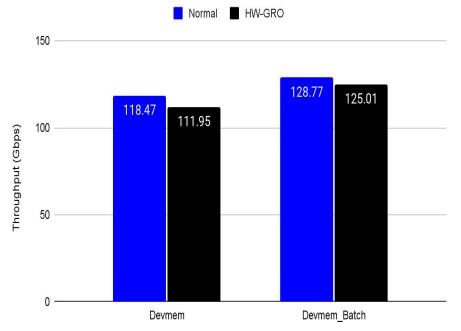
--- System Call Analysis ---Number of recymsg calls: 253592 Number of setsockopt calls: 315547 --- Frags Analysis ---Average frags per recvmsg: 874.644 Minimum frags received: 1 Maximum frags received: 1499 Number of times accumulated frags > 1024: 111714 Number of recymsg with frags > 1024: 97418 Number of recymsg with frags >= 512 && < 1024: 102476 Number of recymsg with frags < 512: 53698 Number of setsockopt avoided (accumulated up to 512): 42242 Percentage >1024/recvmsg: 38.42% Percentage >1024/acc(>1024): 87.20% Percentage <512/recvmsg: 21.17%

On average we are getting more than 512 pages per recvmsg

#### Throughput Devmem - (Batch vs No Batch)



#### Throughput Devmem - 4k MTU - 2 CPUs



### Syscall Costs Before Batching

:14342 (14342), 4745666 events, 100.0%								
syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)	
 recvmsg setsockopt ioctl epoll_wait	615359 530116 1224901 1835		58831.493 22841.143 1232.993 628.492	0.001 0.001 0.001 0.001 0.000	0.112 0.043 0.001 0.343	0.371 0.284 0.604 100.162	0.03% 0.03% 0.09% 28.36%	

#### Syscall Costs After Batching

4139 (4139), 59	904400 event	s, 100.0	0%				2 CPUs
syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
recvmsg	809214	2595		0.001	0.094	0.272	0.03%
setsockopt	530128	0	14209.108	0.001	0.027	0.214	0.03%
ioctl	1609592	0	1544.356	0.001	0.001	0.643	0.07%
epoll_wait	2599	0	725.770	0.001	0.279	100.171	28.20%

## Syscall Costs Before Batching

:3525 (3525), 18	17743 event	s, 100.	0%				1 CPU
syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
setsockopt	294394	0	60365.027	0.001	0.205	324.928	0.98%
recvmsg	206570	4885	30241.574	0.001	0.146	155.632	0.54%
ioctl	403339	0	1305.559	0.001	0.003	21.414	6.40%
epoll_wait	4889	0	1095.391	0.001	0.224	100.170	19.81%

#### Syscall Costs After Batching

cpzc2 (6557), 3	564035 even	ts, 100.	.0%				1 CPU
syscall	calls	errors	total (msec)	min (msec)	avg (msec)	max (msec)	stddev (%)
recvmsg	486578	24534 5	55429.204	0.001	0.114	148.957	0.37%
setsockopt	347744	0 2	27035.294	0.001	0.078	210.768	1.68%
epoll_wait	24541	0	3486.672	0.000	0.142	100.172	5.29%
ioctl	923799	0	2447.129	0.001	0.003	26.322	2.67%

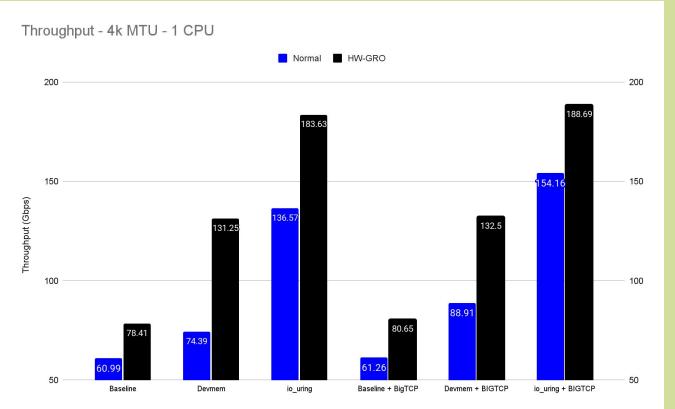
# Perf Devmem (Batch x No Batch)

Devmem - 2 CPUs - Batch (128.77 Gbps)	日 Devmem - 2 CPUs (118.47 Gbps)				
Samples: 199K of event 'cpu-cycles', Event count (approx.): 179161519297	Samples: 199K of event 'cpu-cycles', Event count (approx.): 168906025843				
Overhead Command Shared Object Symbol	Overhead Command Shared Object Symbol				
+ 19,18% :57661 [kernel.vmlinux] [k] tcp_recvmsg_dmabuf	+ 18,01% :18102 [kernel.kallsyms] [k] tcp_recvmsg_dmabuf				
+ 6,18% swapper [idpf] [k] idpf_vport_splitq_napi_poll	+ 8,54% swapper [idpf] [k]idpf_vport_splitq_napi_poll				
+ 3,97% swapper [kernel.vmlinux] [k] napi_pp_put_page	+ 0,29% :18102 [kernel.kallsyms] [k] napi_pp_put_page				
+ 3,86% :57661 [kernel.vmlinux] [k] put_cmsg	+ 5,71% swapper [kernel.kallsyms] [k] tcp_gro_receive				
+ 3,42% swapper [kernel.vmlinux] [k] tcp_gro_receive	+ 5,21% swapper [kernel.kallsyms] [k] napi_pp_ut_page				
+ 3,32% ksoftirqd/0 [idpf] [k] idpf_vport_splitq_napi_poll	+ 3,63% :18102 [kernel.kallsyms] [k] put_cmsg				
+ 3,03% :57661 [kernel,vmlinux] [k] xas_store	+ 2,93% :18102 [kernel.kallsyms] [k] xas_store				
+ 3,00% :57661 [kernel.vmlinux] [k] page_pool_put_netmem_bulk	+ 2,12% :18102 [kernel.kallsyms] [k] xas_find_marked				
+ 2,22% :57661 [kernel.vmlinux] [k] xas_find_marked	+ 1,82% :18102 [kernel.kallsyms] [k] xas_load				
+ 2,05% ksoftirqd/0 [kernel.vmlinux] [k] napi_pp_put_page	+ 1,68% :18102 [kernel.kallsyms] [k] _raw_spin_lock_bh				
+ 1,92% :57661 [kernel.vmlinux] [k] xas_load	+ 1,53% swapper [kernel.kallsyms] [k] napi_build_skb				
+ 1,87% ksoftirqd/0 [kernel.vmlinux] [k] tcp_gro_receive	+ 1,32% swapper [kernel.kallsyms] [k] tcp_v4_rcv				
+ 1,48% :57661 [kernel.vmlinux] [k] tcp_recvmsg_locked	+ 1,31% swapper [kernel.kallsyms] [k] dev_gro_receive				
+ 1,15% :57661 [kernel.vmlinux] [k] tcp_rcv_established	+ 1,26% swapper [kernel.kallsyms] [k] skb_release_data				
+ 1,10% swapper [kernel.vmlinux] [k] napi_build_skb	+ 1,24% :18102 [kernel.kallsyms] [k] tcp_recvmsg_locked				
+ 1,07% :57661 [kernel.vmlinux] [k] xas_clear_mark	+ 1,12% :18102 [kernel.kallsyms] [k] xas_clear_mark				
+ 0,98% swapper [kernel.vmlinux] [k] dev_gro_receive	+ 1,11% :18102 [kernel.kallsyms] [k] tcp_rcv_established				
+ 0,90% swapper [kernel.vmlinux] [k] skb_gro_receive	+ 1,09% swapper [kernel.kallsyms] [k] skb_gro_receive				
+ 0,90% swapper [kernel.vmlinux] [k] skb_release_data	+ 0,87% swapper [kernel.kallsyms] [k]inet_lookup_established				
+ 0,90% swapper [kernel.vmlinux] [k]inet_lookup_established	+ 0,85% :18102 [kernel.kallsyms] [k] tcp_rcv_space_adjust				
+ 0,90% :57661 [kernel.vmlinux] [k] sock_rfree	+ 0,84% swapper [kernel.kallsyms] [k]napi_build_skb				
+ 0,86% :57661 [kernel.vmlinux] [k] skb_try_coalesce	+ 0,82% swapper [kernel.kallsyms] [k] intel_idle				
+ 0,80% :57661 [kernel.vmlinux] [k] tcp_v4_do_rcv	+ 0,82% :18102 [kernel.kallsyms] [k] skb_try_coalesce				
$\overline{T}$ ip: Save output of perf stat using: perf stat record <target workload=""></target>	Fip: If you have debuginfo enabled, try: perf report -s sym,srcline				
Devmem - 1 CPU - Batch (92.42 Gbps)					
Samples: 100K of event 'cpu-cycles', Event count (approx.): 92873468822	Samples: 100K of event 'cpu-cycles', Event count (approx.): 93207866858				

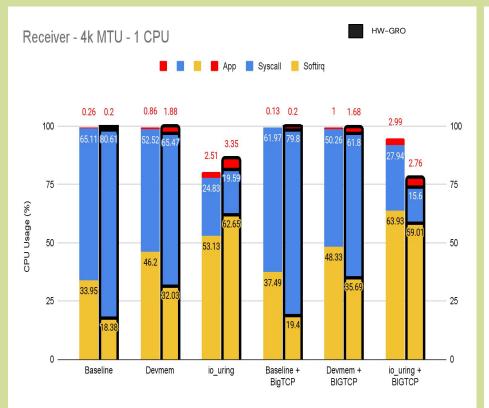
Devmem - 1 CPU - Batch (92.42 Gbps)			H Devmem - 1 CPU (74.39 Gbps)					
		Samples: 100K of event 'cpu-cycles', Event count (approx.): 93207866858						
Overhead Command Shared Object	Symbol	01		Command	Shared Obje		Symbol	
+ 8,32% :52790 [kernel.vmlinux]	[k] tcp_recvmsg_dmabuf	+		:8076			[k] gen_pool_has_addr	
+ 8,03% :52790 [idpf]	[k] idpf_vport_splitq_napi_poll	+		:8076	[kernel.vml	inux]	[k] gen_pool_free_owner	
+ 5,66% :52790 [kernel.vmlinux]	[k] tcp_gro_receive	+		:8076	[idpf]		[k] idpf_vport_splitq_napi_poll	
+ 5,33% :52790 [kernel.vmlinux]	[k] put_cmsg	+		:8076			[k] tcp_recvmsg_dmabuf	
+ 4,36% :52790 [kernel.vmlinux]	[k] xas_store	+		:8076	[kernel.vml	inux]	[k] tcp_gro_receive	
+ 3,23% :52790 [kernel.vmlinux]	[k] napi_pp_put_page	+		:8076			[k] gen_pool_alloc_algo_owner	
+ 2,98% :52790 [kernel.vmlinux]	[k] xas_find_marked	+		:8076			[k] put_cmsg	
+ 2,97% :52790 [kernel.vmlinux]	[k] gen_pool_free_owner	+		:8076				
+ 2,94% :52790 [kernel.vmlinux]	[k] gen_pool_has_addr	+					[k] napi_pp_put_page	
+ 2,85% :52790 [kernel.vmlinux]	[k] xas_load	+		:8076	[kernel.vml	inux]	[k] xas_find_marked	
+ 2,10% :52790 [kernel.vmlinux]	[k] gen_pool_alloc_algo_owner	+		:8076			[k] xas_load	
+ 1,85% :52790 [kernel.vmlinux]	[k] check_preemption_disabled	+		:8076			[k] _raw_spin_lock_bh	
+ 1,62% :52790 [kernel.vmlinux]	[k] dev_gro_receive						[k] check_preemption_disabled	
+ 1,50% :52790 [kernel.vmlinux]	[k] xas_clear_mark						[k] dev_gro_receive	
+ 1,42% :52790 [kernel.vmlinux]	[k] page_pool_put_netmem_bulk							
+ 1,36% :52790 [kernel.vmlinux]	<pre>[k] page_pool_alloc_frag_netmem</pre>						[k] net_is_devmem_iov	
+ 1,32% :52790 [kernel.vmlinux]	[k] net_is_devmem_iov	+					[k] page_pool_alloc_frag_netmem	
+ 1,23% ksoftirqd/0 [idpf]	[k] idpf_vport_splitq_napi_poll	+		:8876			[k] inet_gro_receive	
+ 1,08% :52790 [kernel.vmlinux]	[k] skb_gro_receive	+					[k] skb_gro_receive	
+ 1,06% :52790 [kernel.vmlinux]	[k]xa_cmpxchg	+		:8076			[k]napi_build_skb	
+ 1,02% :52790 [kernel.vmlinux]	[k]napi_build_skb	+		:8076			[k] eth_type_trans	
+ 0,97% :52790 [kernel.vmlinux]	[k] inet_gro_receive						[k] page_pool_put_unrefed_netmem	
+ 0,96% :52790 [kernel.vmlinux]	[k] eth_type_trans	+	0,74%	:8076	[kernel.vml	inux]	[k] napi_build_skb	

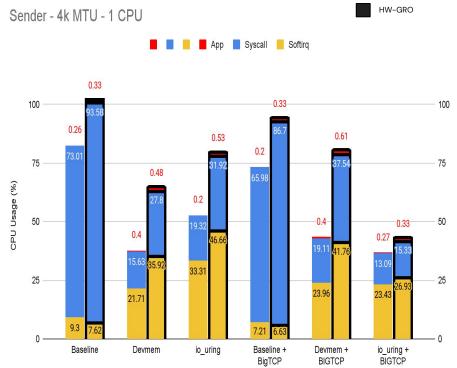
#### Results Putting app + softirq on one core

# Throughput - 1 CPU

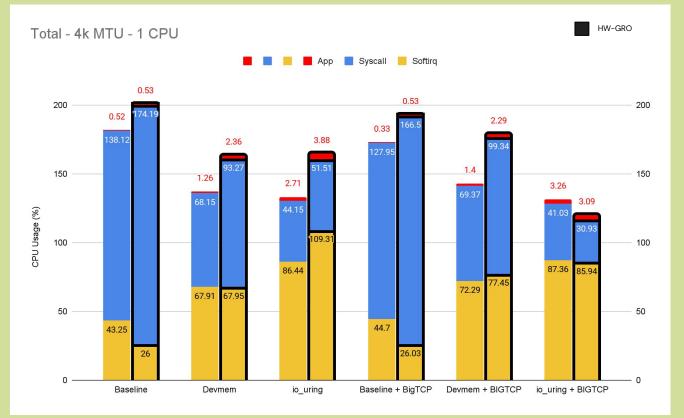


#### CPU Usage - 1 CPU - Receiver vs Sender



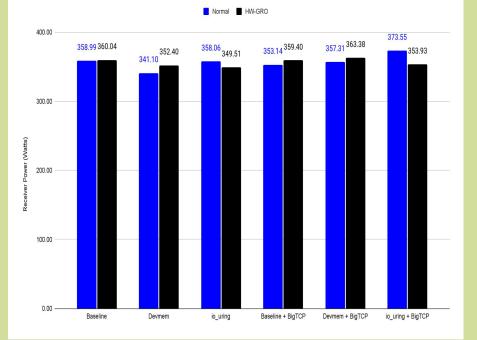


# CPU Usage - 1 CPU

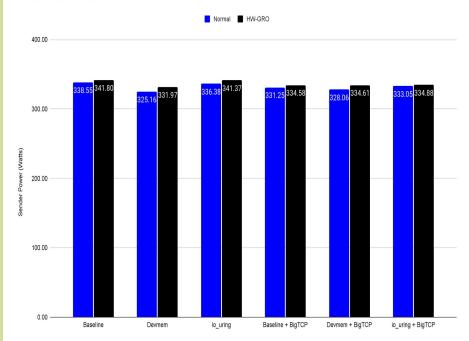


# Power -1 CPU

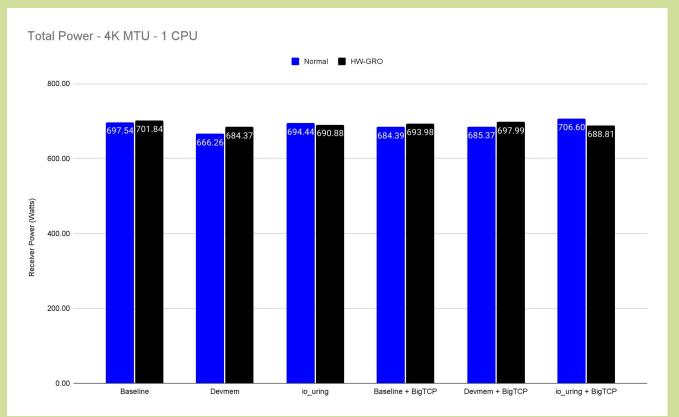
#### Receiver Power - 4K MTU - 1 CPU



Sender Power - 4K MTU - 1 CPU



# Power Total - 1 CPU



# **ROI** Performance Metrics

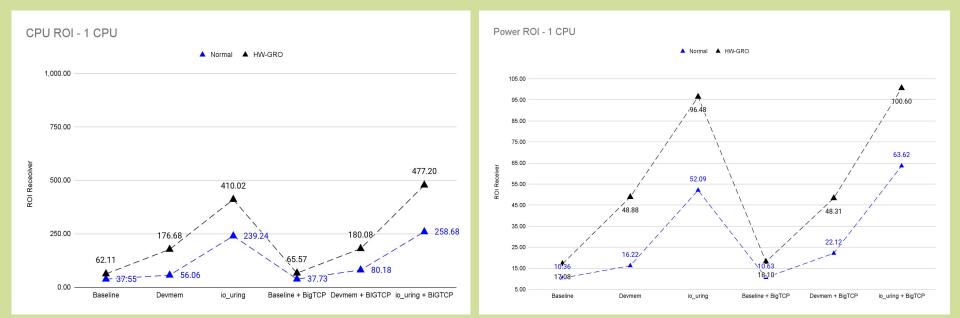
When comparing different results for <u>throughput</u>, it is often hard to say which results gives you the best return on investmen(ROI). We came up with two ROI formulas:

 $\frac{T^n}{\sum C_i} \, \overset{\text{T=Throughput achieved in Gbp/s. n=2}}{\subset \text{-compute cost to achieve the throughput. Summed across all CPUs}} \, \mathcal{T}^n$ 

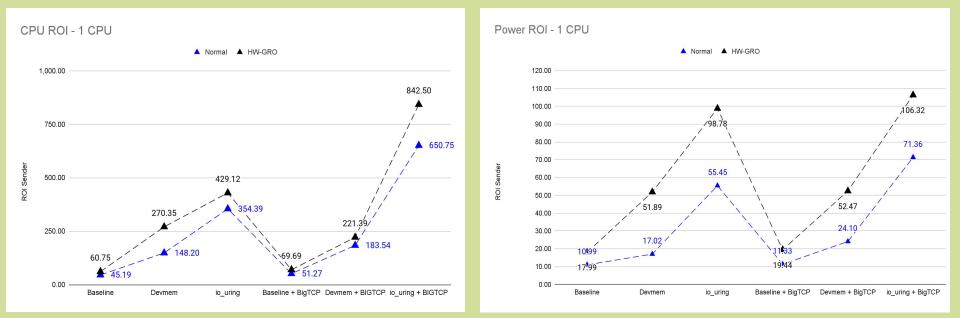
 $T^n$  T=Throughput achieved in Gbp/s. n=2  $\overline{P^m}$  P=Power consumed by in watts. m=1

We pick n=2 to emphasize throughput as the goal. So 10 gbps using 200% cpu is considered to be better ROI than 5 gbps using 100% cpu

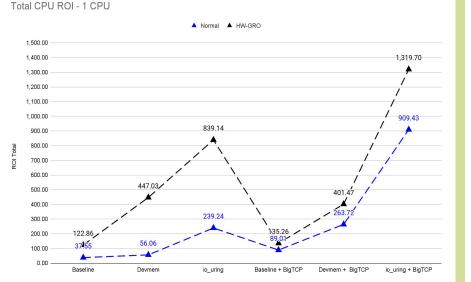
## ROI Receiver - 1 CPU

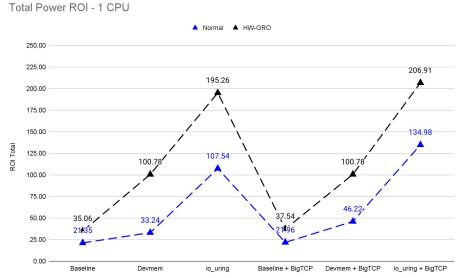


## ROI Sender - 1 CPU



# ROI Total - 1 CPU





# **Conclusions And Recommendations**

- The two ZC interfaces are still fluid but nevertheless deliver on the message
- Devmem is syscall heavy while io\_uring is kernel/softirq heavy
- io\_uring demonstrates a clear advantage over devmem, however we need to test scaling to much higher rates (>200Gbps)
- Page pool API or perhaps driver use of the PP API could be improved
- HW GRO clearly provides a big advantage

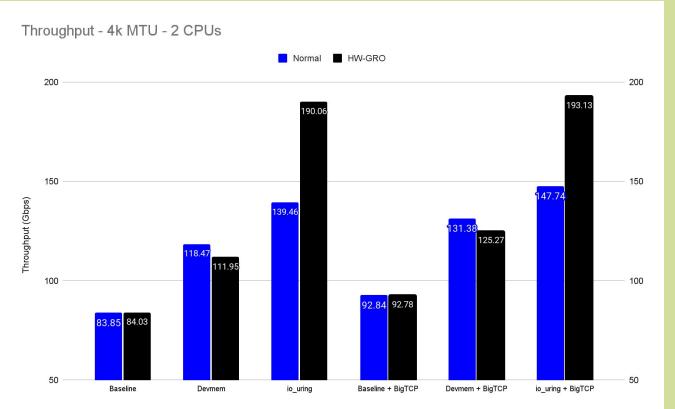
# Acknowledgements

We could not have done it without their help and patience. Appreciated!

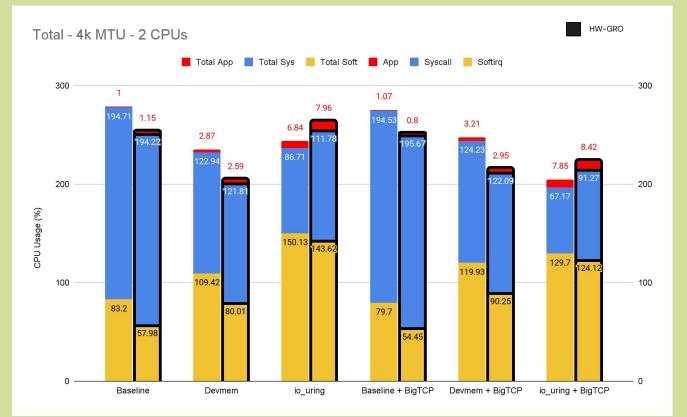
- Anjali S. Jain Instrumental for anything Intel!
- Sridhar Samudrala for driver patches
- Mina Almasry for Devmem and guidance into gpud/nccl integration
- Pavel Bengukov for io\_uring

#### **Back Slides**

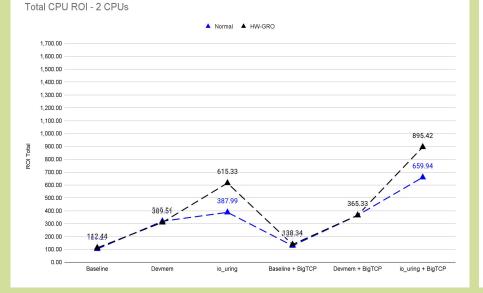
# Throughput - 2 CPUs

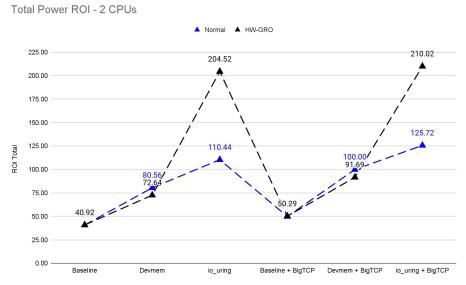


# CPU Usage - 2 CPUs

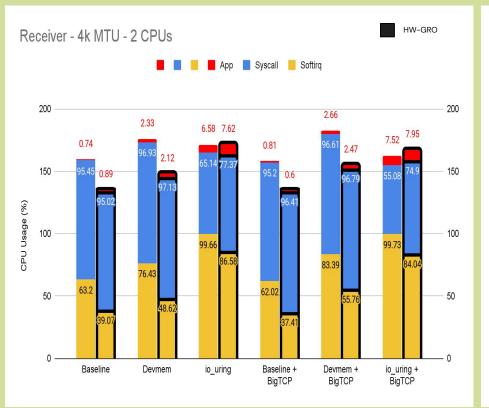


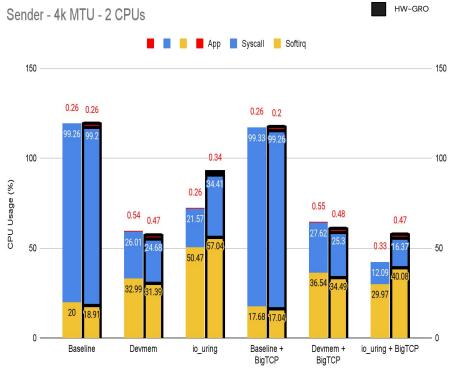
## ROI Total - 2 CPUs





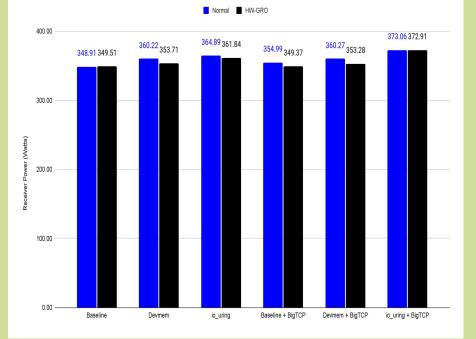
#### CPU Usage - 2 CPUs - Receiver vs Sender



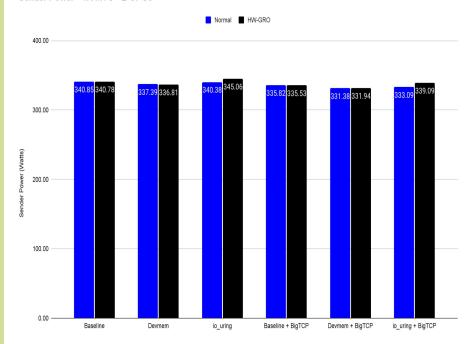


# Power - 2 CPUs

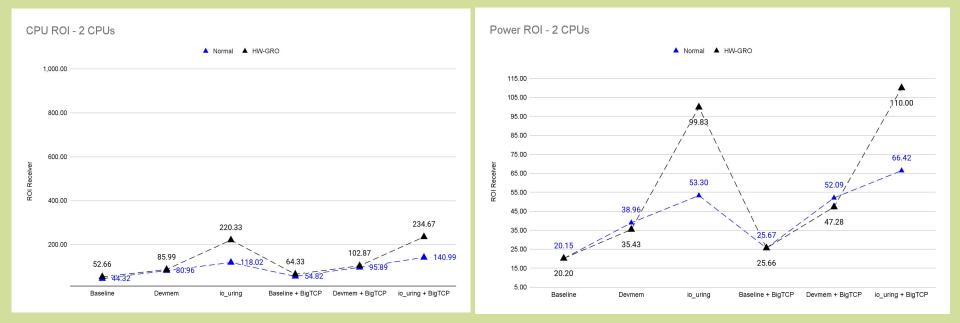
#### Receiver Power - 4K MTU - 2 CPUs



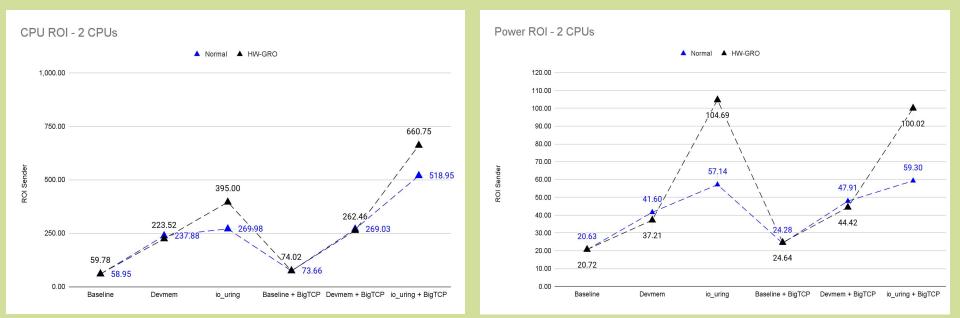
Sender Power - 4K MTU - 2 CPUs



#### ROI Receiver - 2 CPUs



## ROI Sender - 2 CPUs

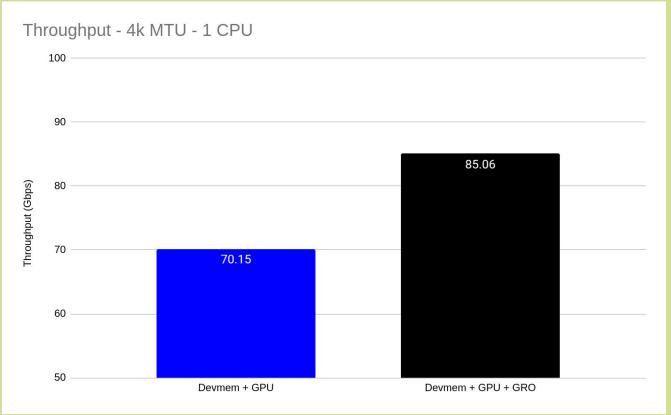


# Sample Accelerator

Cheap GPU: NVidia T1000

- Sitting on a PCIE3 x16 riser card (max expected ~100Gbps)
- 4G RAM
- Run some dummy CUDA code to keep the GPU awake

# Throughput - I CPU - GPU Mem



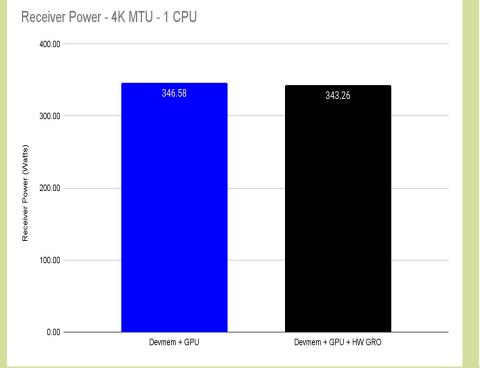
# Nvtop - Receiver

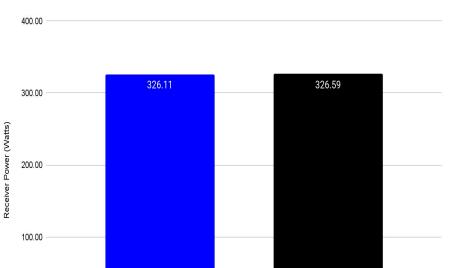


#### CPU Usage - 1 CPU - GPU Mem

Receiver - 4k MTU - 1 CPU 📕 SUT app 📕 SUT sys 📒 SUT soft 100 13.98 75 CPU Usage (%) 54.5 50 25 28.09 Devmem + GPU Devmem + GPU + HW GRO Sender - 4k MTU - 1 CPU 📕 TG app Normal 📒 TG sys Normal 📒 TG soft Normal 50 0 40 0 30 CPU Usage (%) 27.57 20 17.39 10 Devmem + GPU Devmem + GPU + HW GRO

# Power - 1 CPU - GPU Mem





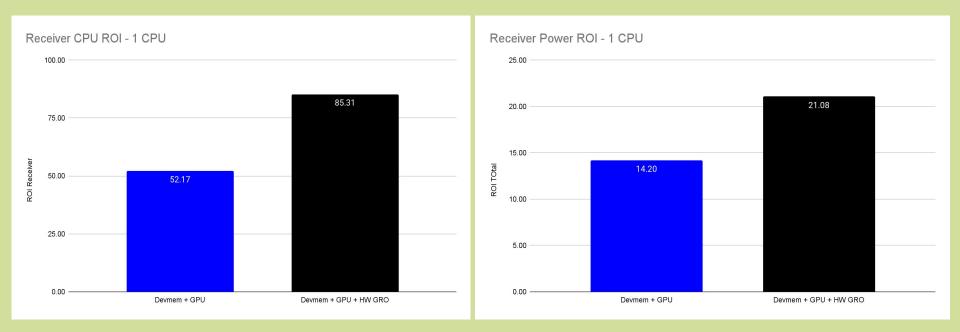
Devmem + GPU + HW GRO

Devmem + GPU

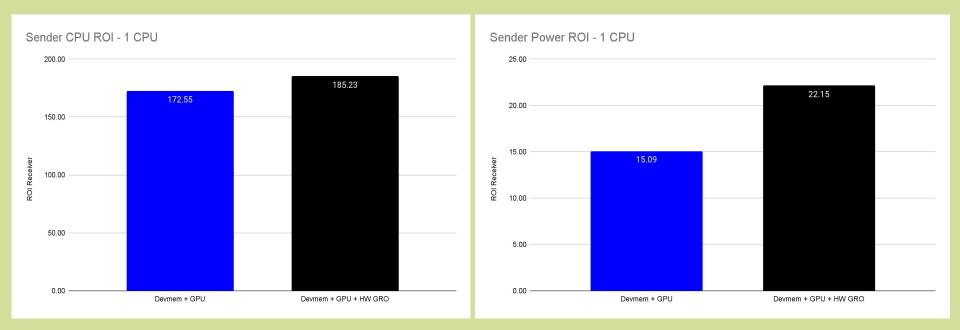
Sender Power - 4K MTU - 1 CPU

0.00

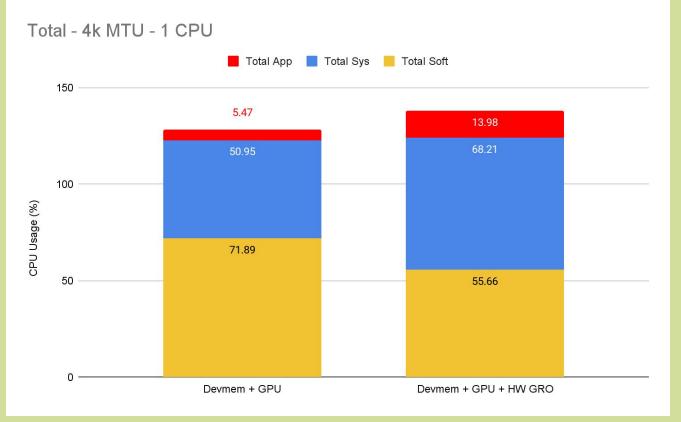
## ROI Receiver - 1 CPU - GPU Mem



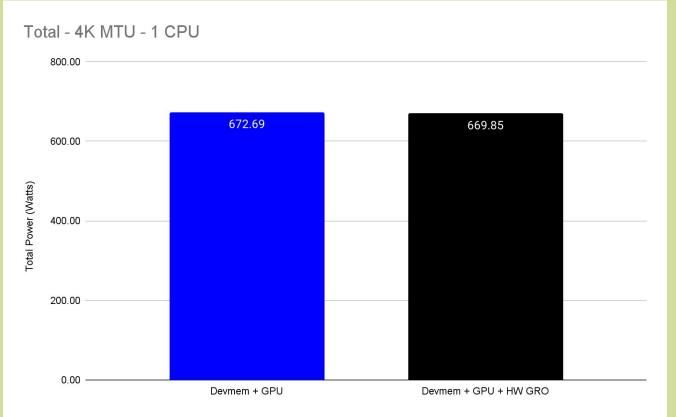
## ROI Sender - 1 CPU - GPU Mem



# CPU Usage Total - 1 CPU - GPU Mem



# Power Total - 1 CPU - GPU Mem



# ROI Total - 1 CPU - GPU Mem

