



DPDK

DATA PLANE DEVELOPMENT KIT

DPDK PMD for AF_XDP

Zhang Qi, Li Xiaoyun

Intel

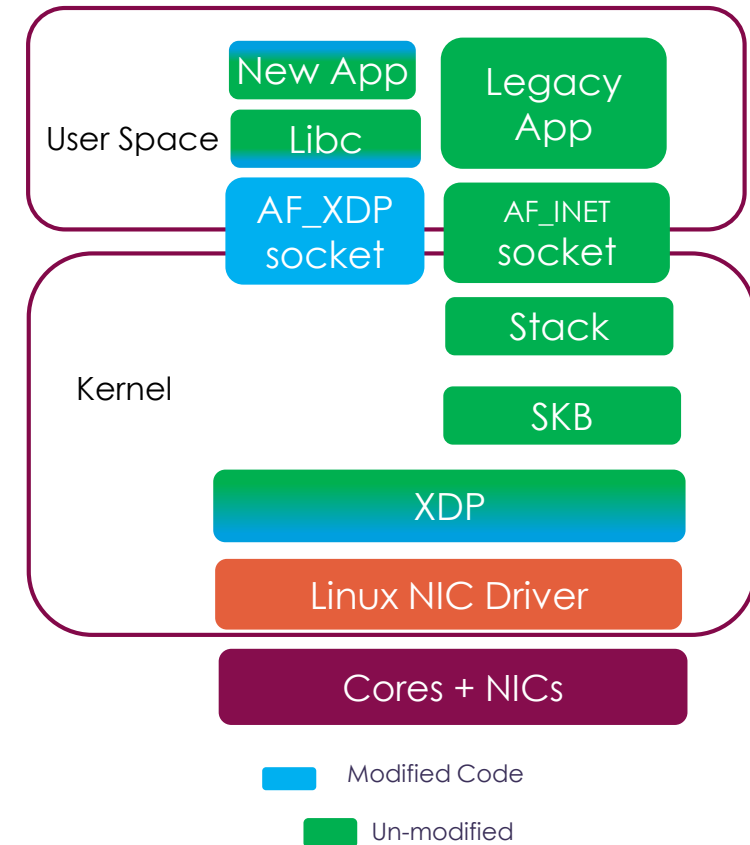
DPDK UserSpace - Dublin 2018

Agenda

- ❑ Introduction to AF_XDP
- ❑ DPDK PMD for AF_XDP
- ❑ Future Work

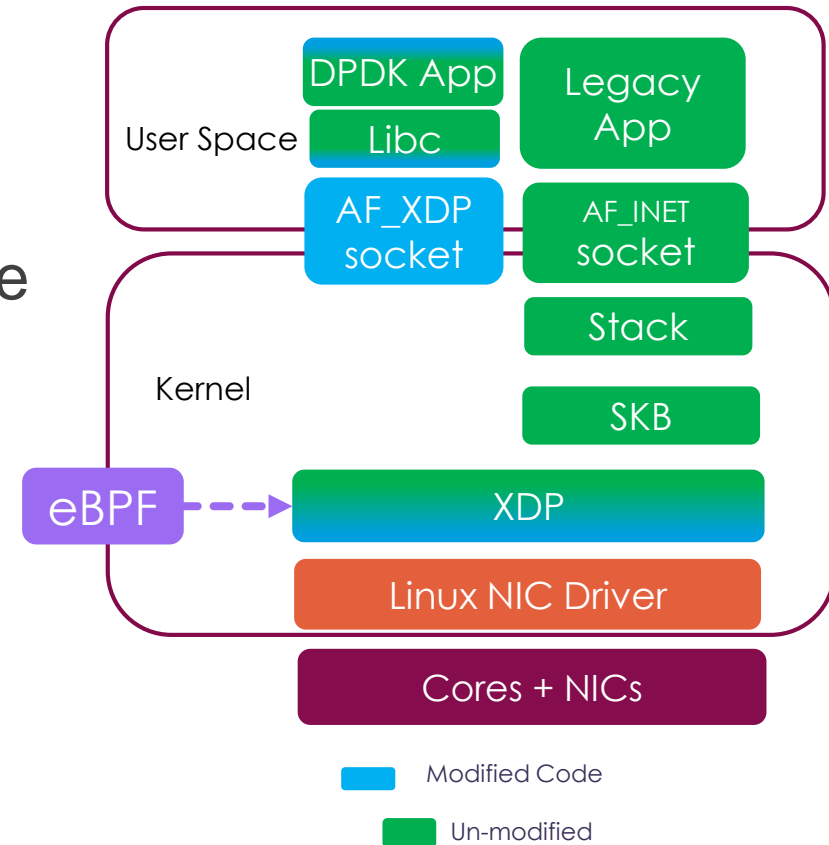
Introduction to AF_XDP

- An upgraded version of AF_PACKET
 - Use XDP program to trigger Rx path for selected queue
- XDP programs can redirect frames to a memory buffer in user-space by eBPF.
- DMA transfers use **user space memory** (Zero Copy)



Introduction to AF_XDP -- Benefits

- Performance improvement
 - Zero copy between user space and kernel space
 - Achieve 3-20X times improvement comparing to AF_PACKET
- Connect the XDP pass-through to user-space directly
 - An eBPF program that processes packets can be forwarded to an application in a very efficient way
- For DPDK
 - No change to DPDK apps, kernel driver handles hardware
 - Provide a new option for users



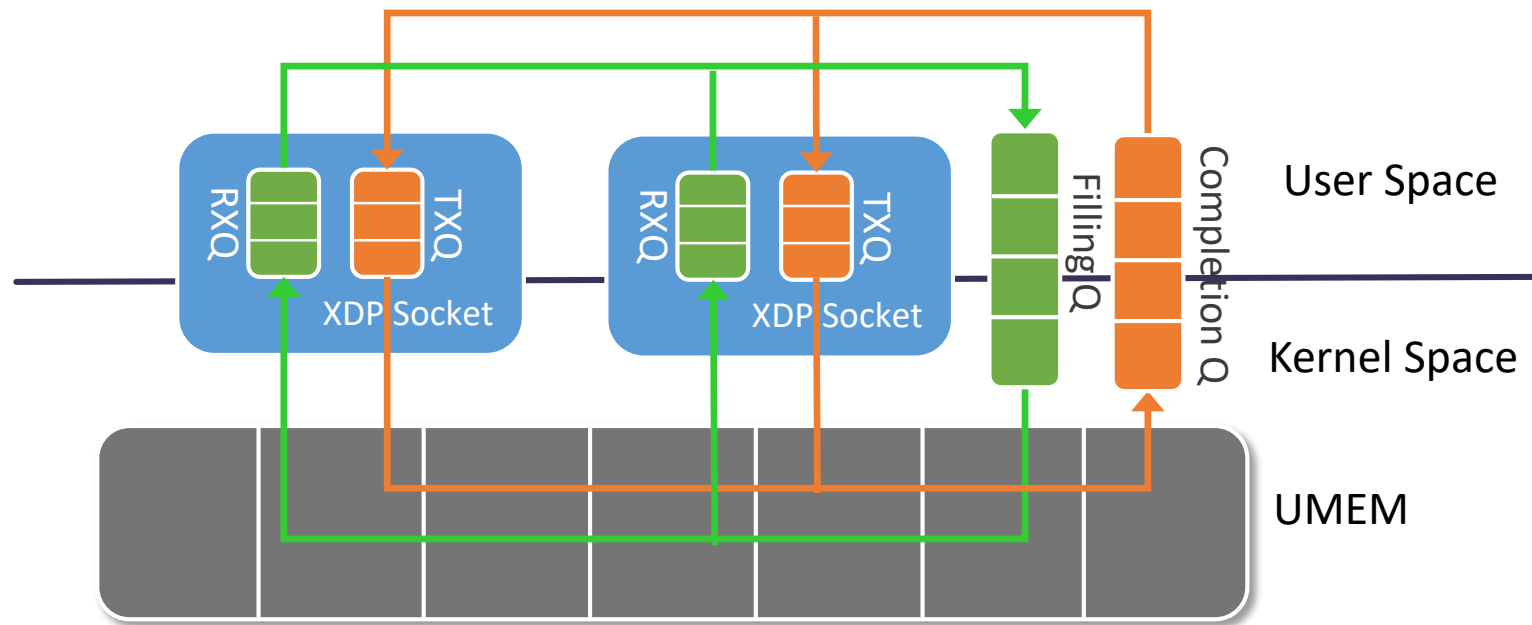
Introduction to AF_XDP

- AF_XDP socket

```
sfd = socket(PF_XDP, SOCK_RAW, 0);
bufs = calloc(num_bufs, FRAME_SIZE);
setsockopt(sfd, SOL_XDP, XDP_MEM_REG, bufs);
setsockopt(sfd, SOL_XDP, XDP_{RX|TX|FILL|COMPLETE}_RING,
ring_size);
mmap(..., sfd, .....); /* map kernel rings */
bind(sfd, "/dev/eth0", queue_id, .....);
for (;;) {
    read_process_send_messages(sfd);
};
```

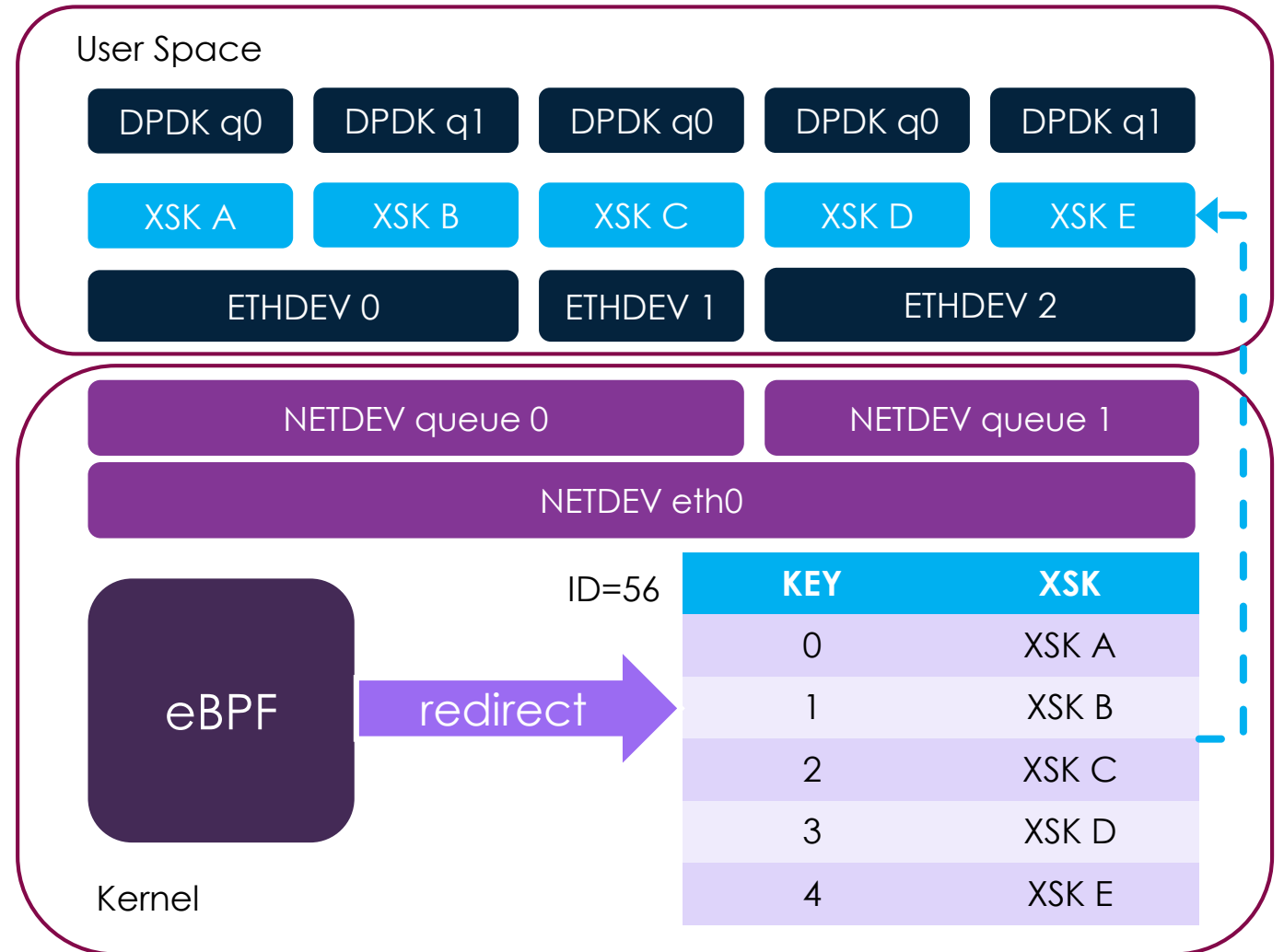
Introduction to AF_XDP

- Datapath



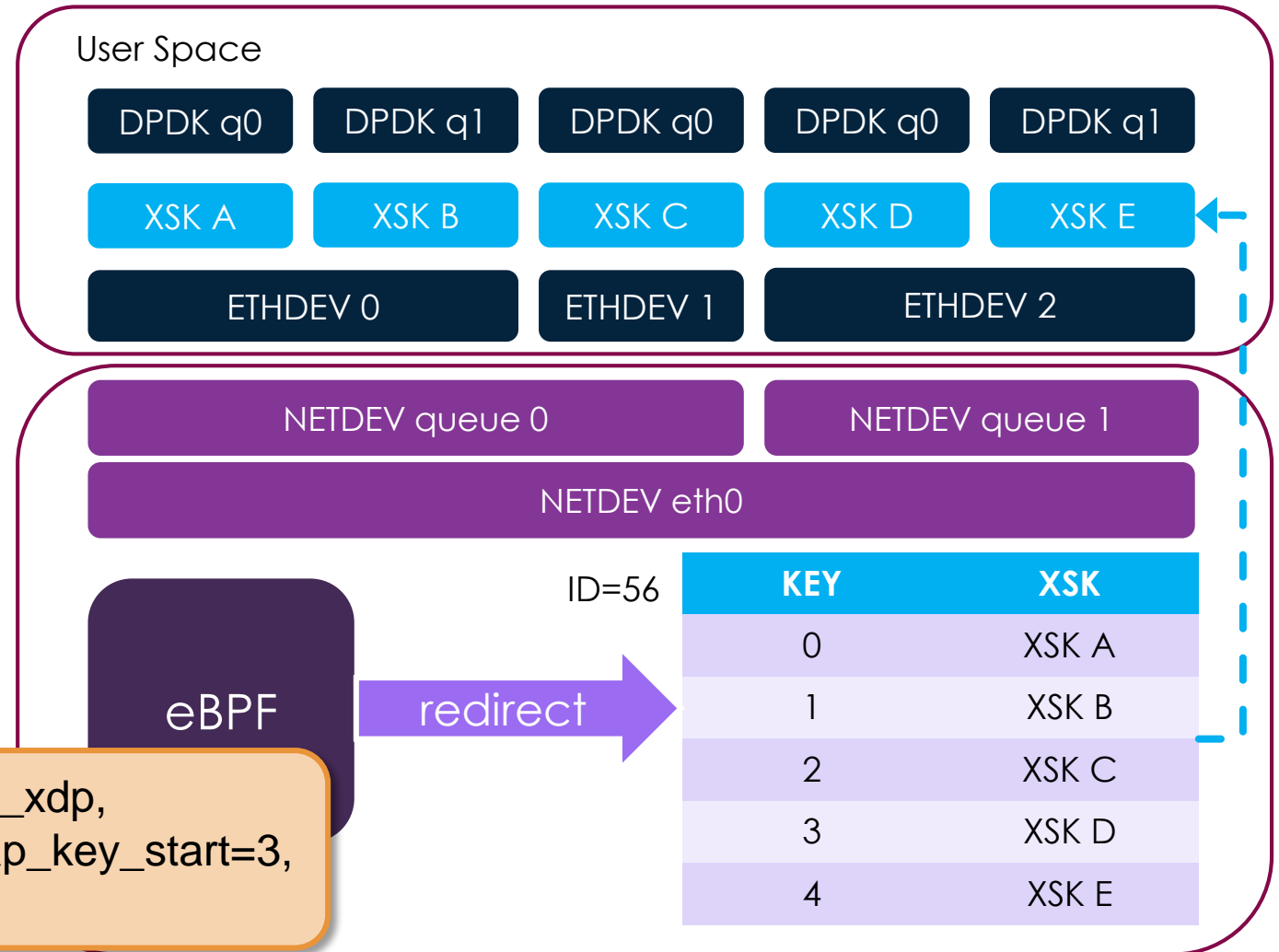
DPDK PMD for AF_XDP

- Enable AF_XDP socket



DPDK PMD for AF_XDP

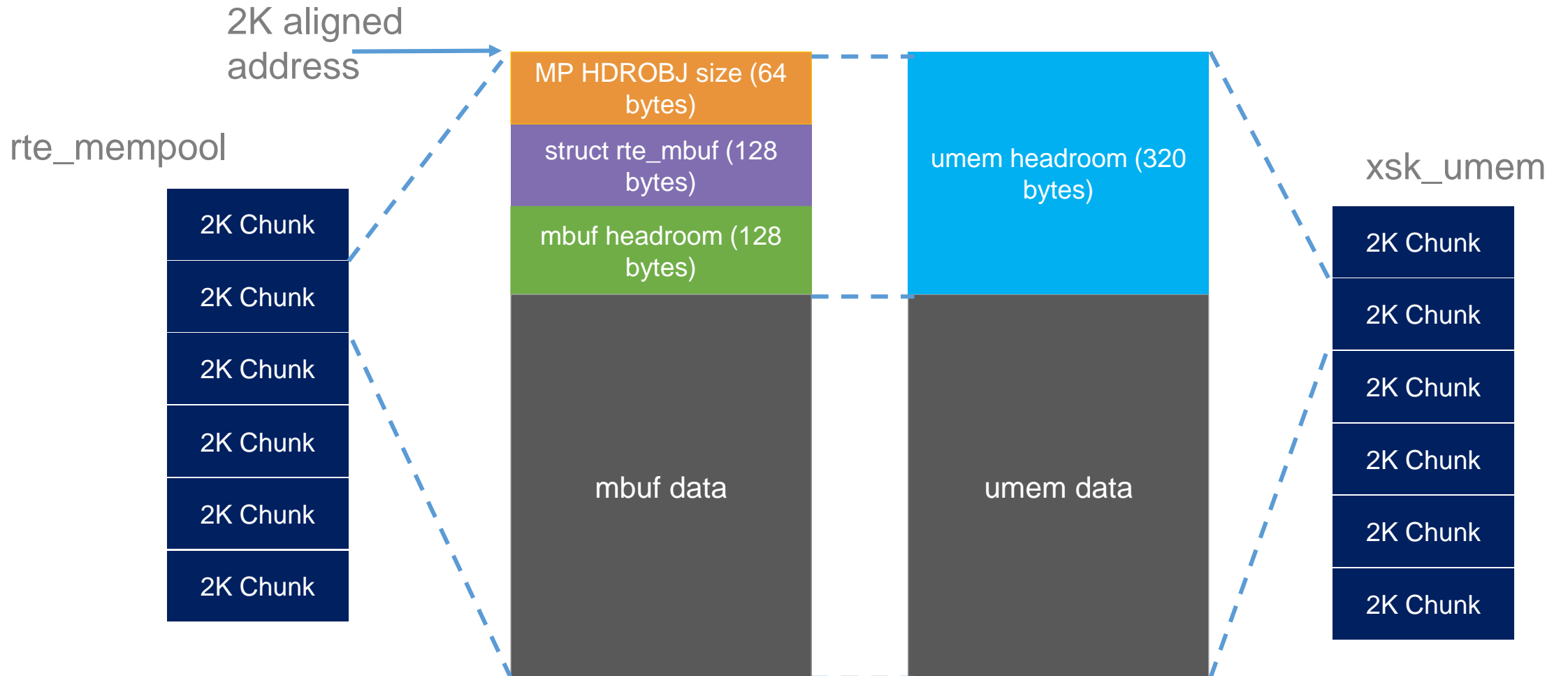
- Load eBPF program
- Link ebpf program to specific network interface
- Expose the xdp socket redirect map id and entries number to user



```
./build/app/testpmd -c 0xc -n 4 --vdev eth_af_xdp,  
iface=en0,queue=1,xsk_map_id=56,xsk_map_key_start=3,  
xsk_map_key_count=2 -- -i --rxq=2 --txq=2
```

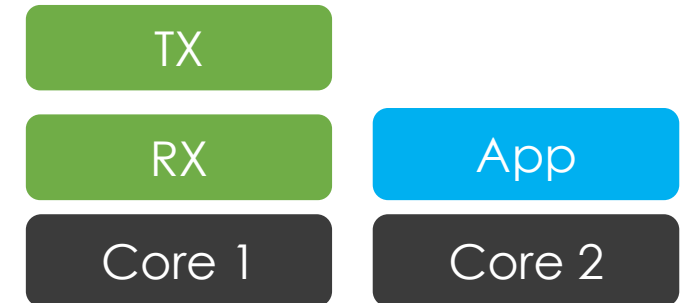

DPDK PMD for AF_XDP

- Mapping between `rte_mempool` and `xsk_umem`



AF_XDP Experimental Setup

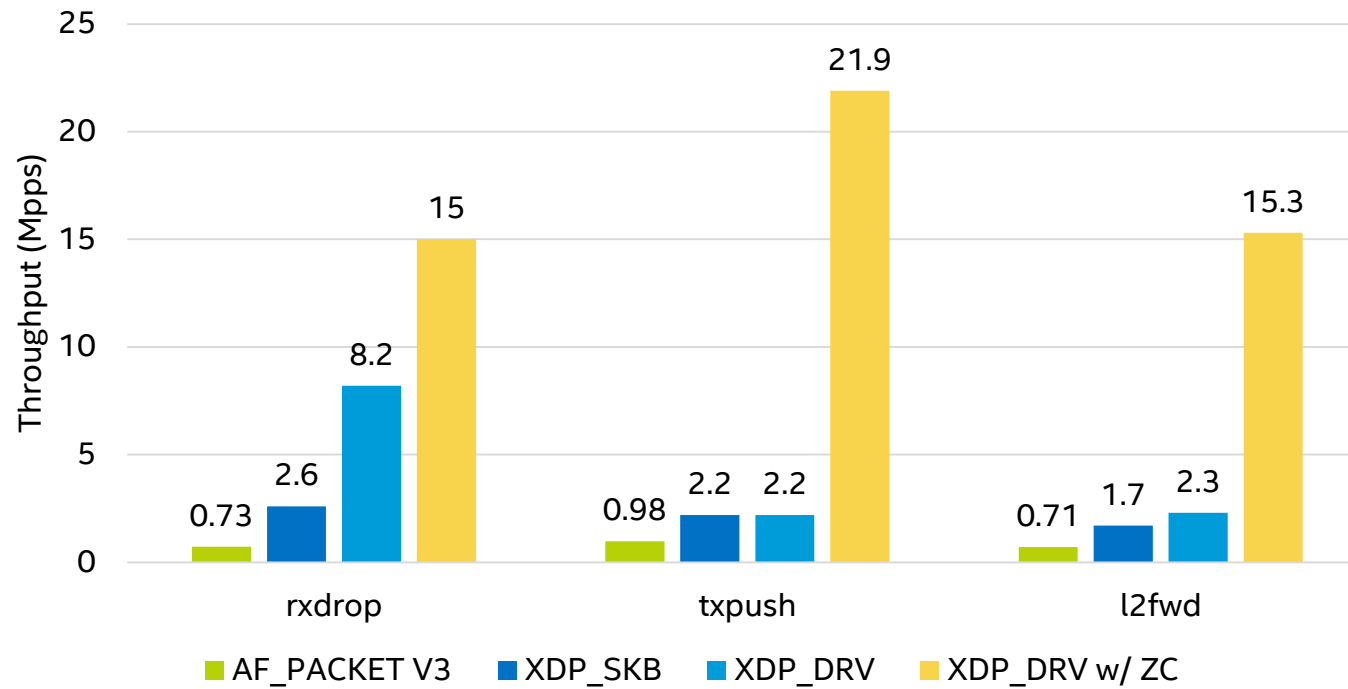
- V2 patch set of AF_XDP published on April 27, 2018 + ZC patches in house
 - Very few optimizations performed
- Broadwell E5 2660 @ 2.0 GHz
- IXIA load generator blasting at full 40 Gbit/s
- Intel XL710 card (40G, i40e driver)



<http://lists.openwall.net/netdev/2018/08/28/62>

AF_XDP performance

- PMD performance is aligned with this



<http://lists.openwall.net/netdev/2018/08/28/62>

- Refine zero copy enabling in PMD
- Usertools for AF_XDP in DPDK
- Follow up kernel AF_XDP's changes
- Performance optimization

Q & A

<http://mails.dpdk.org/archives/dev/2018-August/109791.html>

Zhang Qi

qi.z.zhang@intel.com

Li Xiaoyun

xiaoyun.li@intel.com