External buffer

Raslan Darawsheh
Mellanox

External buffer

- First was introduced by Olivier in his presentation in 2016.
- The support for External buffer has been there since DPDK 18.05

External buffer: Cont’d

Mbuf with external data buffer

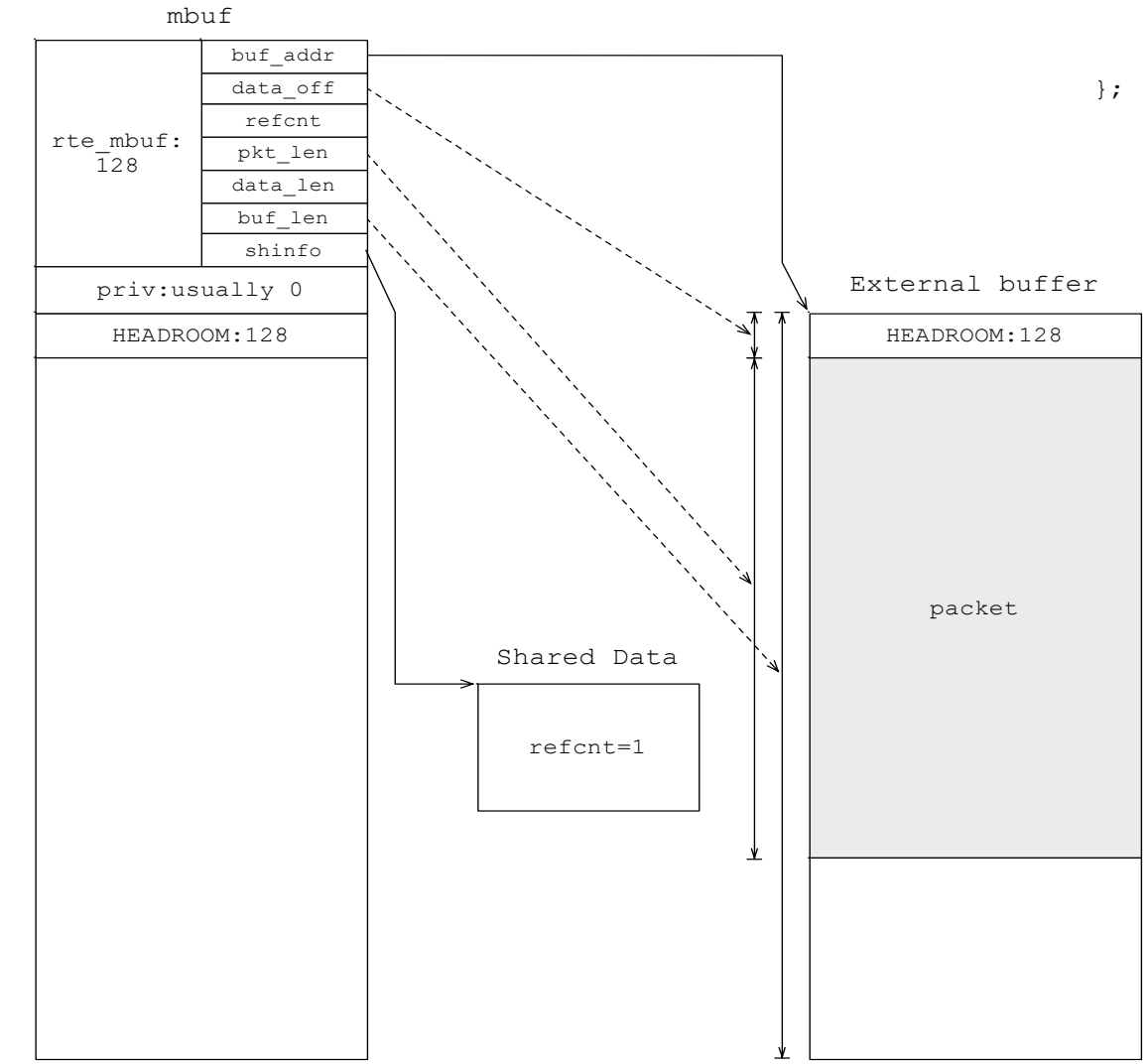
External buffer: Cont'd

Discuss: mbuf with external buffer (1)

- Currently, a mbuf embeds its data (direct), or references another mbuf (indirect)
- It could make sense to have mbuf referencing external memory
- Use cases: virtual drivers, server applications, storage, traffic generators

External buffer: Cont'd

mbuf EXT_ATTACHED_MBUF




```
struct rte_mbuf_ext_shared_info {
      rte_mbuf_extbuf_free_callback_t free_cb; /**< Free callback function */
      void *fcb_opaque;                        /**< Free callback argument */
      rte_atomic16_t refcnt_atomic;            /**< Atomically accessed refcnt */
};
```

- Attach external buffer (non-mbuf) to a mbuf
  - Attached rte_pktmbuf_attach_extbuf()
    - shinfo must be provided
      - User managed buffer
      - rte_pktmbuf_ext_shinfo_init_helper()
        - Helper function to simply spare a few bytes at the end of the buffer for shared data.
    - m->ol_flags |= EXT_ATTACHED_MBUF;
    - rte_pktmbuf_reset_headroom() is needed to reset HEADROOM
  - Detached by rte_pktmbuf_detach_extbuf()
    - Same as rte_pktmbuf_detach()
    - rte_mbuf_ext_refcnt_update(m->shinfo, -1);
    - Call shinfo->free_cb() if m->shinfo->refcnt gets zero

External Buffer

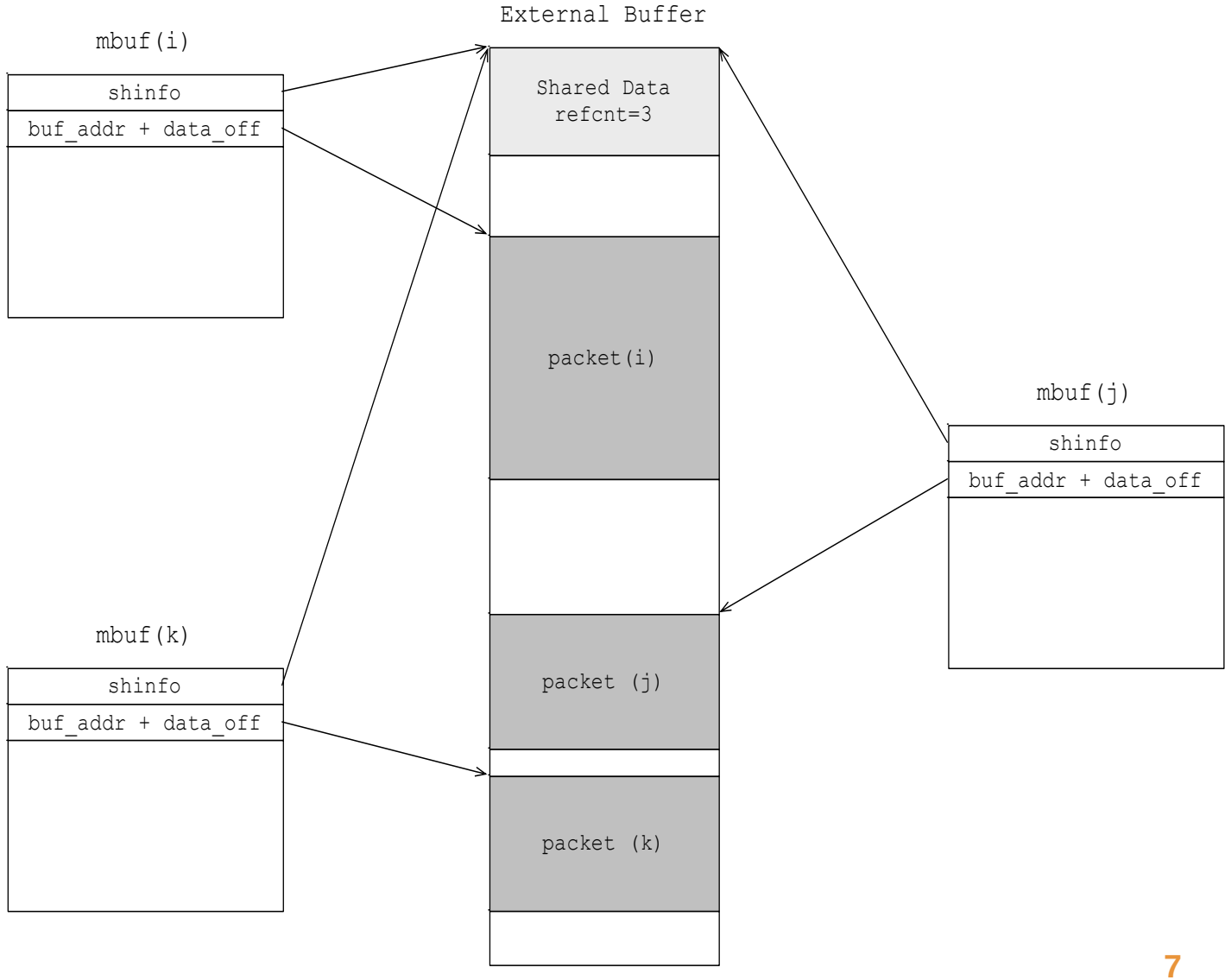
USE CASES

Use cases : #1 Storage applications

Single buffer is shared by multiple Mbufs.

No need to copy data to mbuf for Tx

Common shinfo external buffer is read-only



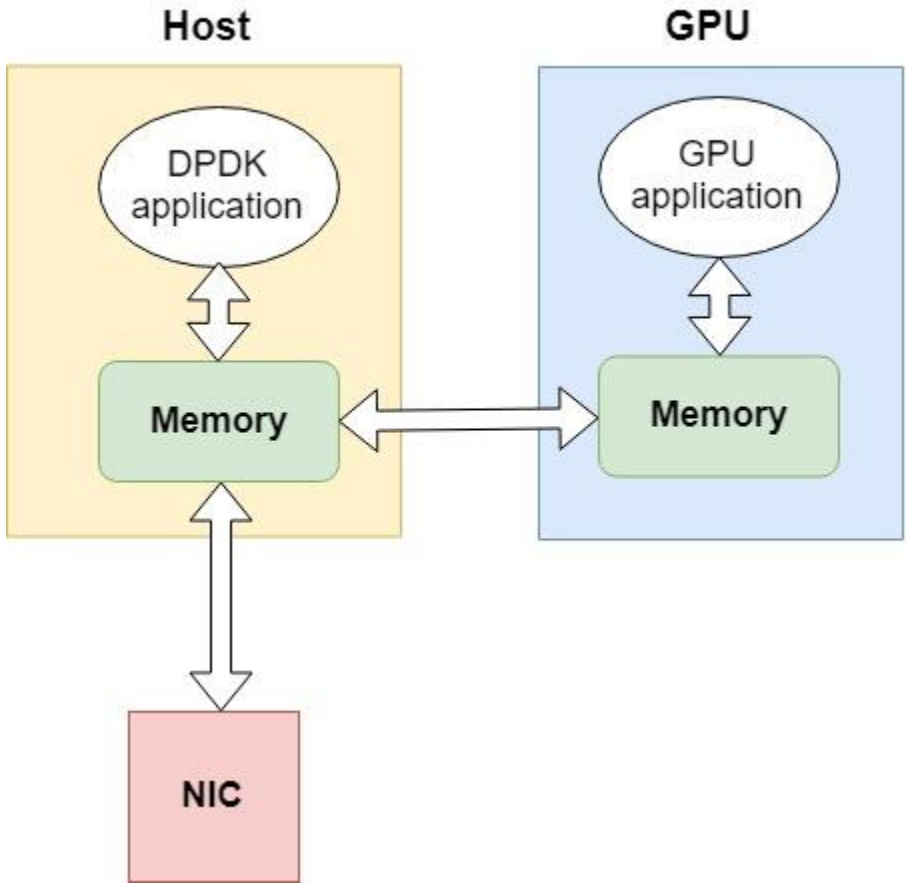
Use cases: #2 GPU

GPUs (Graphics Processing Units) are being used to accelerate complex and time consuming tasks in a range of application.

Typically GPU's don't manage the packet send/receive, only the data processing.

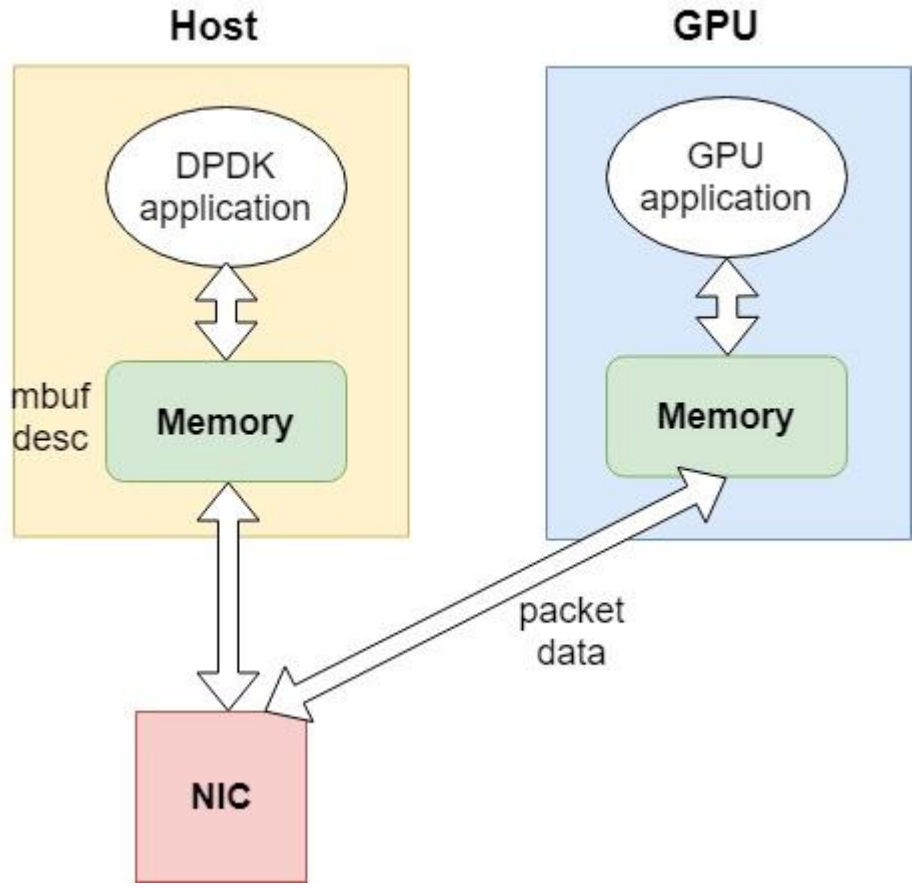
With the current mbuf scheme the GPU must copy in the packet data from the host memory.

The external mbuf enables true zero copy for the GPU and hence improve the performance significantly



Use cases: GPU – Cont'd

The Mempool on host memory is populated w/ mbuf descriptor allocated from the host memory and it has some external buffer attached to it.



DPDK
DATA PLANE DEVELOPMENT KIT

GPU use case

Some testpmd POC Result

	Default behavior	With External buffer
Single Core	Rx-pps: 5532097 ~ 5.5 Mpps.	Rx-pps: 10869938 ~ **10.8 Mpps.**
	Tx-pps: 5539771 ~ 5.5 Mpps.	Tx-pps: 10869938 ~ **10.8 Mpps.**
	CPU cycles/packet=213	**CPU cycles/packet=98 (217% improvement).**
4 Cores	**Default behavior**	**With External buffer**
	RX-pps: 11852669 ~ 11.8 Mpps.	Rx-pps: 38054341 ~ **38.05 Mpps.**
	Tx-pps: 11852669 ~ 11.8 Mpps.	Tx-pps: 38054341 ~ **38.05 Mpps.**
	CPU cycles/packet=910	**CPU cycles/packet=160 (568% Improvement).**

DPDK
DATA PLANE DEVELOPMENT KIT

Future plan

- IND_ATTACHED_MBUF or EXT_ATTACHED_MBUF is set in m->ol_flags even though it is not an offload.

- Open question:

- Would It be better to have an explicit way to say an mbuf is read only ? Currently, it is implicit. mbuf/ext_buf having refcnt > 1 .

DPDK
DATA PLANE DEVELOPMENT KIT

“

Thank you all

”