Mellanox Bifurcated DPDK PMD

Rony Efraim

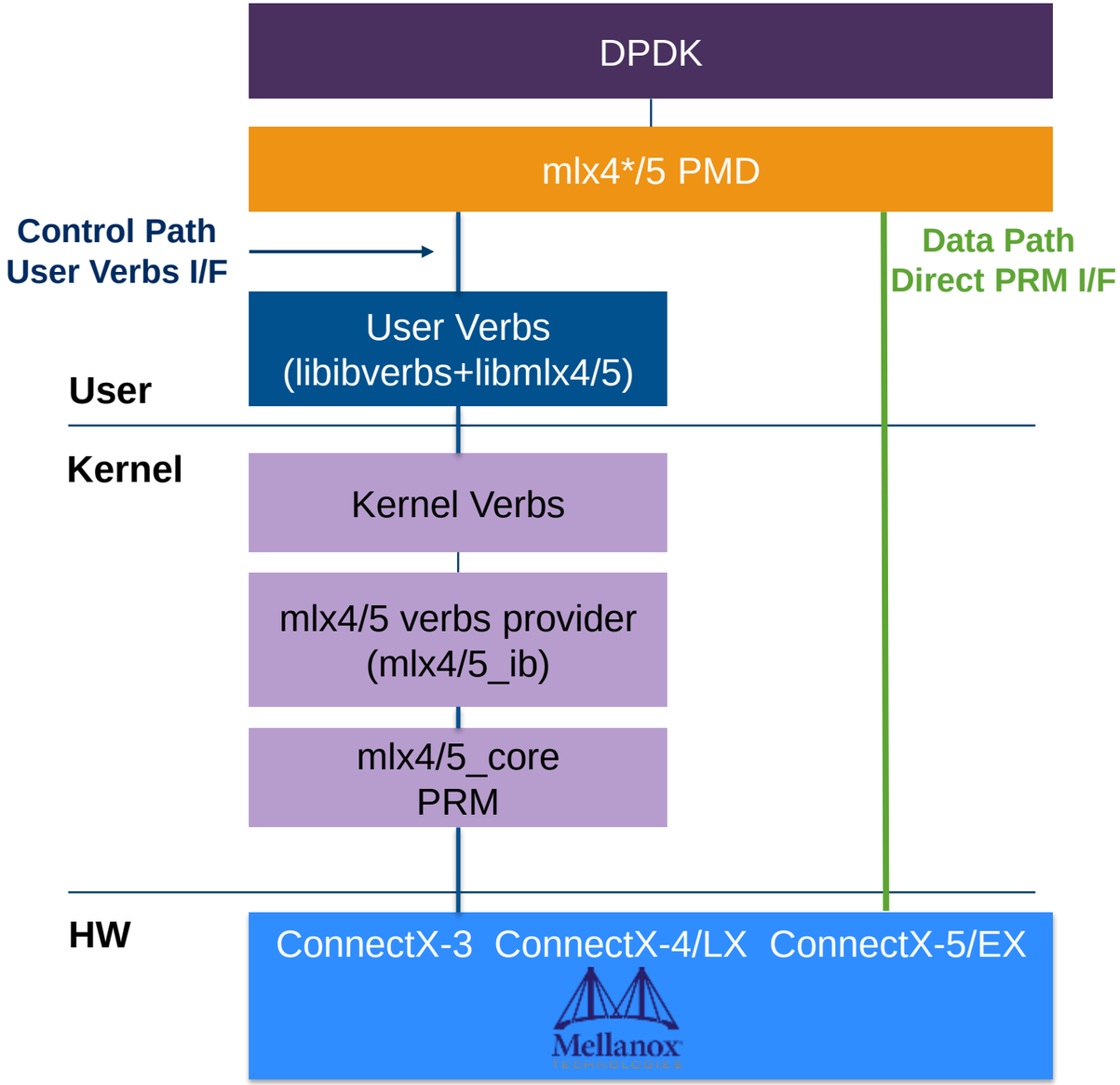
DPDK summit Nov 2017

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™

Mellanox PMD

- Mellanox PMD is built on top of **libibverbs** using the Raw Ethernet Accelerated Verbs API.
- Verbs is a standard API of openfabrics.org
- The PMD directly accesses the NIC for fast data path
- Control path is going through the kernel modules.
- Unlike other PMDs, Mellanox PMD is yet another user space application and the NIC is still being controlled by the kernel



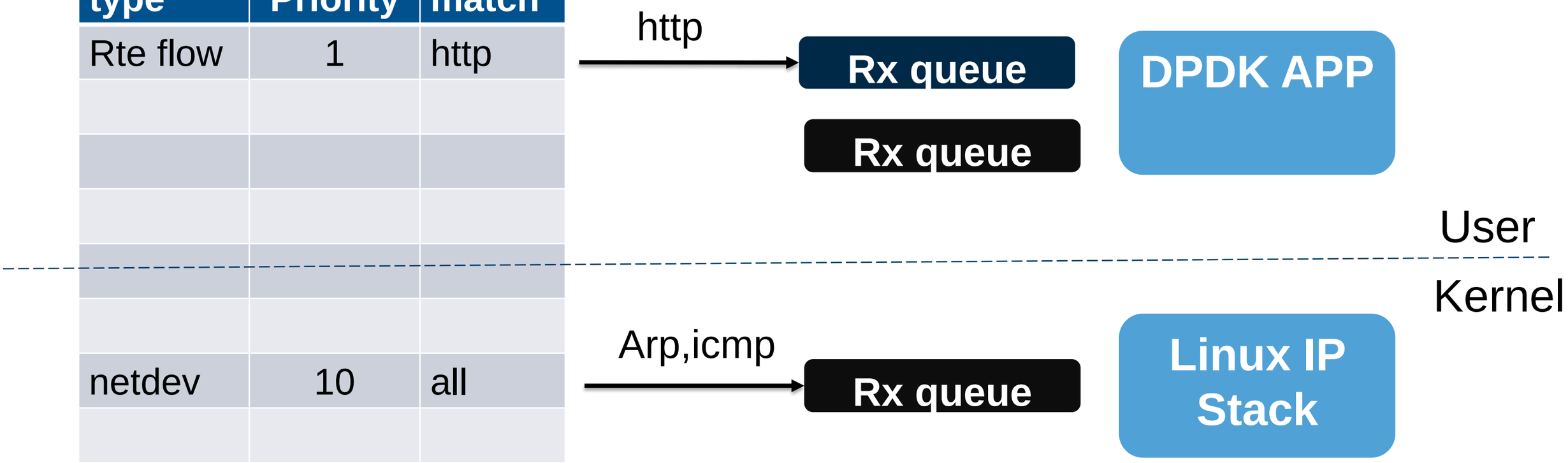
Kernel driver

- Control commands (e.g ip link, ethtool) will work as usual
- Mellanox PMD relies on system calls for control operations such as querying/updating the MTU and flow control parameters.
- For security reasons and robustness, the PMD only deals with virtual memory addresses.
- DPDK steers the required traffic by using rte flow.
- All the leftover is going to the kernel driver.
- The kernel handles the leftover traffic and can reply for arp, ping ….

Example

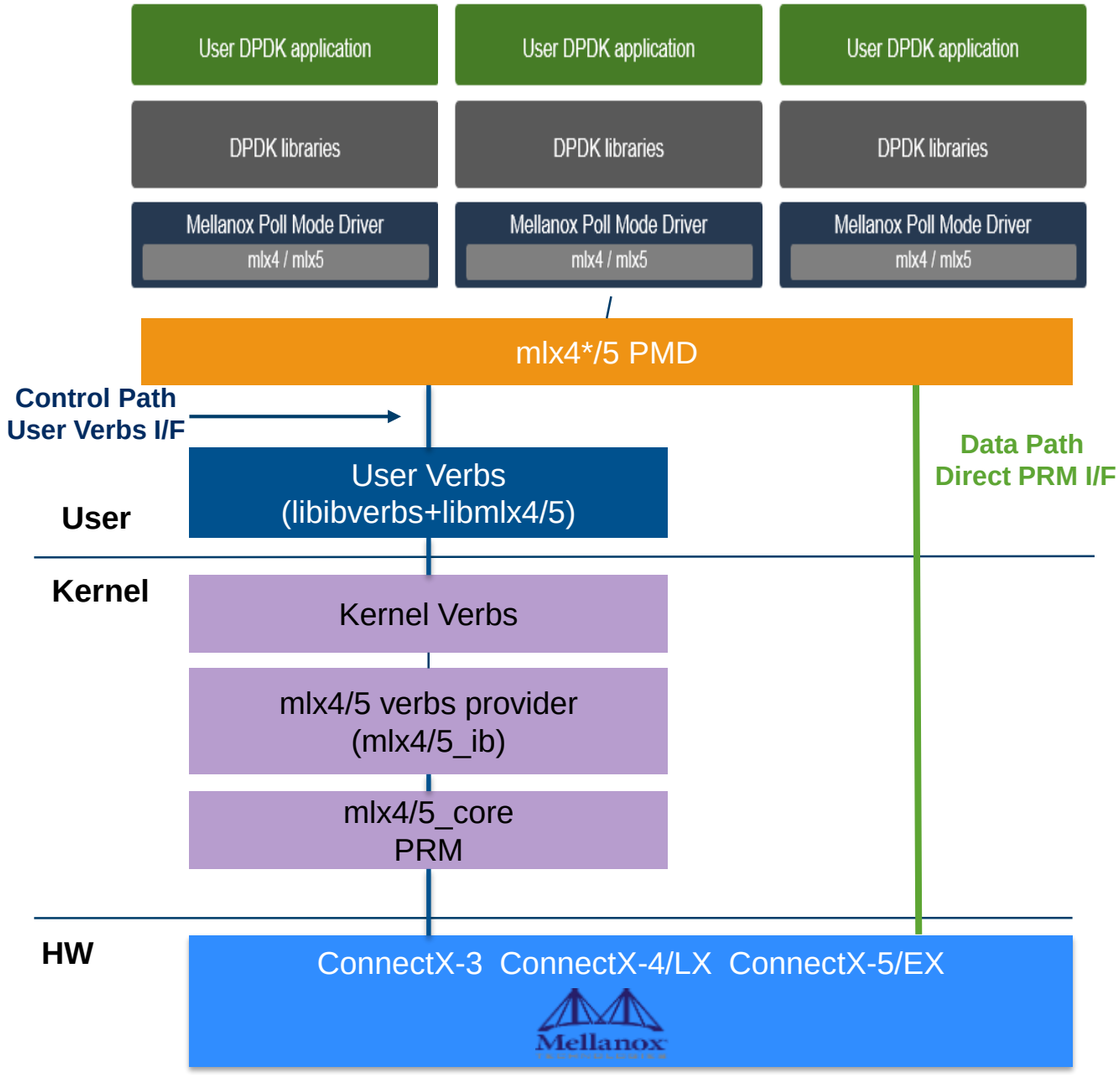
RX packets

type	Priority	match
Rte flow	1	http
netdev	10	all



Multiple DPDK Applications on single device

- You can run multiple applications on a single function.
- resources allocation are handled by the kernel combined with hardware specifications that allow it to handle virtual memory addresses directly ensure that DPDK applications cannot access random physical memory or memory of other processes.
- Steering is done by the application when adding a filter.
- A packet can be steered to multiple queues and/or applications.
- Does not require SR-IOV. Multiple apps work over VF/PF
- Designed to be run as non root thanks to verbs




time for questions

Thank You

Mellanox TECHNOLOGIES

Connect. Accelerate. Outperform.™