PMDLINK - BRIDGING THE GAP BETWEEN DPDK AND HARDWARE PACKET PROCESSING

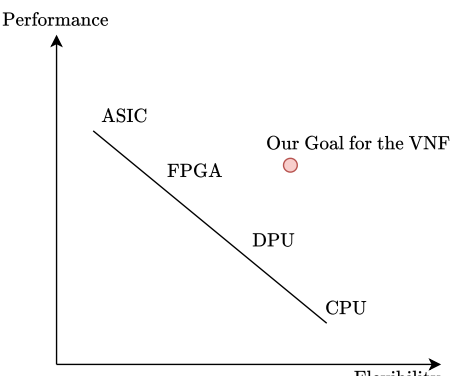
Rubens Figueiredo[1,3], Andreas Kassler[2,3], Hagen Woesner[1], Holger Karl[4], Nic Hart[1]

[1]BISDN GmbH,
[2]Deggendorf Institute of Technology,
[3]Karlstad University,
[4]Hasso-Plattner-Institute

April 4, 2025

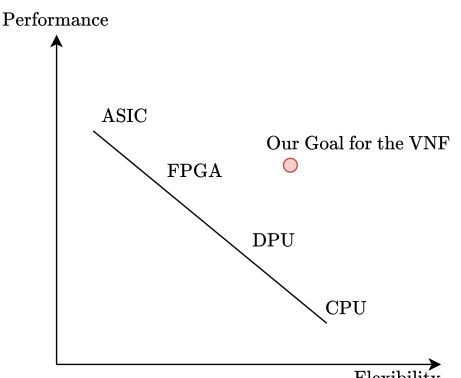
Commodity Hardware And Packet Processsing Offloads

- Software-based packet processing struggles at high speeds

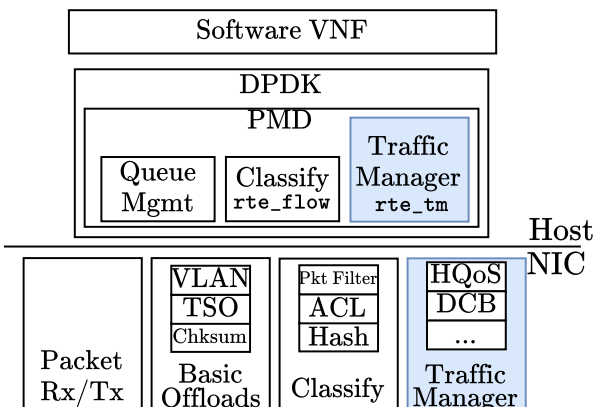


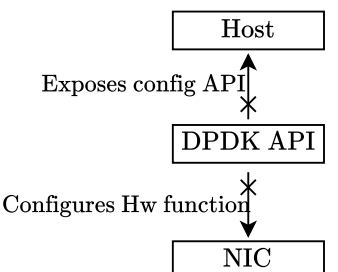
Commodity Hardware And Packet Processsing Offloads

- Software-based packet processing struggles at high speeds
- Hardware offload as candidate to improve performance
  - Programmable ASIC
  - smartNIC
  - **Commodity Hardware**
- How to leverage commodity hardware NIC features to improve performance of the VNF?



DPDK AS INTERFACE TO NIC

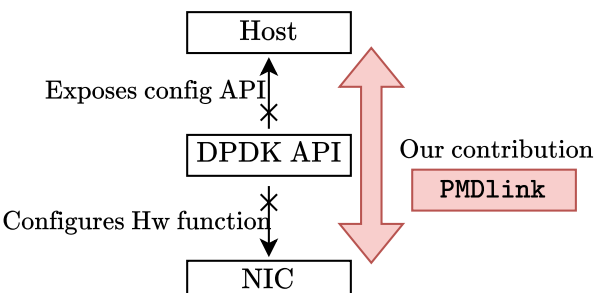




What was our issue? `rte_tm` implementation of `IAVF` Traffic Manager is *restricted*

DPDK as Interface to NIC

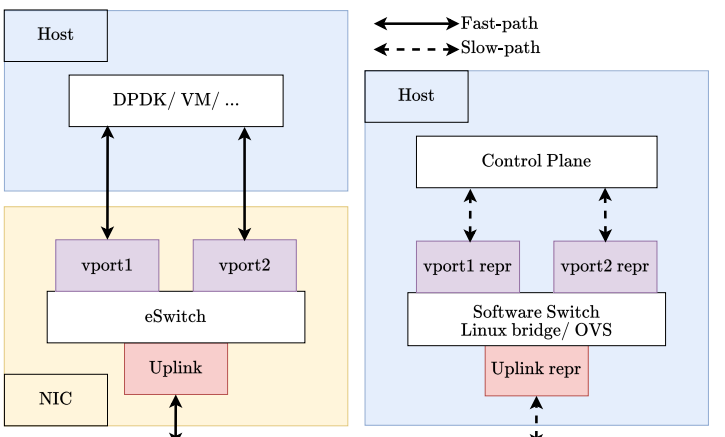
Our approach Direct hardware offload API through `PMDlink`



Addressing the Missing Pieces

Legacy and Switchdev Modes

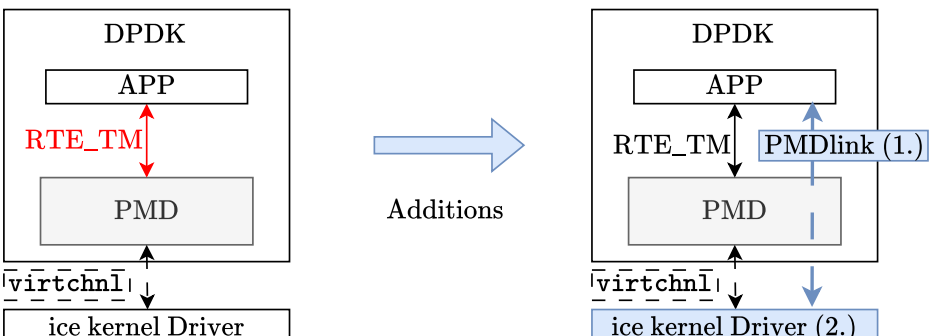
- `eSwitch`: hardware component inside modern NICs
- **Legacy mode**: Limited offloads (MAC/ VLAN)
- **Switchdev**:
  - Advanced offload enabled
  - In Intel E810, *Hierarchical QoS*



Tu, William, Michal Swiatkowski, and Yossi Kuperman. "Shared Memory Pool for Representors.": https://netdevconf.info/0x18/docs/netdev-0x18-paper41-talk-paper.pdf

Addressing the Missing Pieces

Extending DPDK and Intel Drivers



- **Problem** RTE_TM only allows for IEEE 802.1 Qaz (ETS)[1] configuration
- **Additions**
  1. Add `rte_eth_dev_send_vf_msg` and similar method in `IAVF` driver
  2. Modify ice kernel driver to accept HQoS messages.

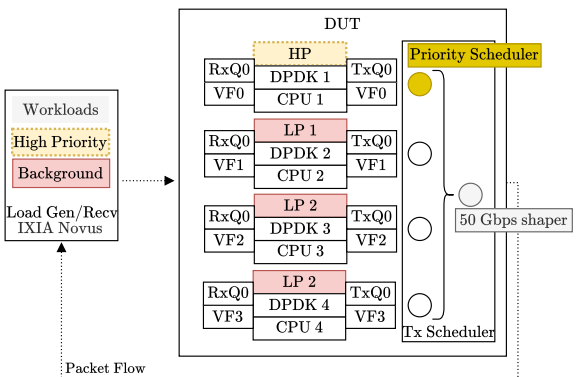
[1] Enhanced Transmission Selection: https://1.ieee802.org/dcb/802-1qaz/

Impact of Background Traffic on High-Priority Latency

Experimental Results

How to achieve high-throughput and low latency packet processing on *commodity hardware*?

Traffic Profile	Details	Packet Size
HP	1 Gbps	178 B
Background	10-95 Gbps	1518 B

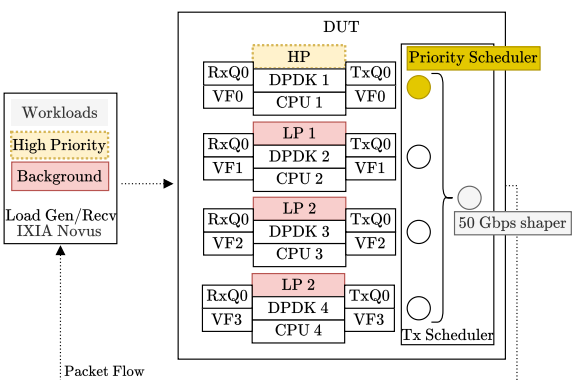


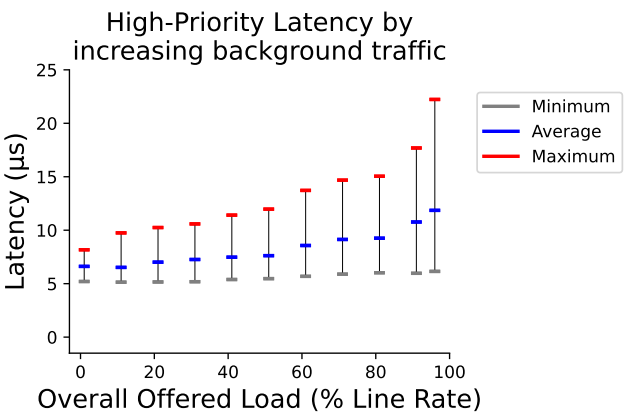
Figueiredo, Rubens, et al. "Quality of Service Performance of Multi-Core Broadband Network Gateways." 2024 8th Network Traffic Measurement and Analysis Conference (TMA). IEEE, 2024.

Impact of Background Traffic on High-Priority Latency

Experimental Results

How to achieve high-throughput and low latency packet processing on *commodity hardware*?





Takeaway: SR-IOV *and* hardware-acceleration for fast packet processing.

Hierarchical QoS executed at hardware speeds without sacrificing flexibility.

Figueiredo, Rubens, et al. "Quality of Service Performance of Multi-Core Broadband Network Gateways." 2024 8th Network Traffic Measurement and Analysis Conference (TMA). IEEE, 2024.

Challenges of Hardware Integration

Future Work

- Expressiveness of existing API implementations
- Other hardware functionality, such as header processing, parsing, matching, ...
- Compatibility with smartNIC, other PMD

Conclusion

For Now

`PMDlink` as a generic API for hardware configuration

Combining

- *Flexibility* of software processing, and
- *Performance* of hardware execution.