Learning from the DPDK

How the other half lives

Stephen Hemminger

May 28, 2018

DPDK Status

- Governed by Linux Foundation
- Active Community
- Contributors mostly HW vendors (Intel, ARM, NIC,...)
- Main OS usage: FD.io and OpenVswitch
- Many Proprietary consumers
- Just added BPF

- Total throughput over 1Tbit/sec
- Vector Packet Processor
- Only uses DPDK drivers
- Batch packet per step
- Performance regression testbed

DPDK Good

- Performance
- Easier to run on ancient kernel
- BSD License = proprietary enabler
- Easier to innovate

- Static QoS
- Fixed ACL's using AVX
- Crytpo HW and SW offload
- Event HW device

- Usability
    - API stability
    - HW catchall
    - NIH
    - Weak config model
- Power Management
- Security

Performance

- Optimize virtio and vhost
- AVX instructions in drivers and firewall
- Pre-computed Longest Prefix Match
- Lockless Tx and Rx path

Recommendations

- Batching everywhere
- Decompose pipeline
- Lockless ring
- Pinning resources to owner (no sharing)