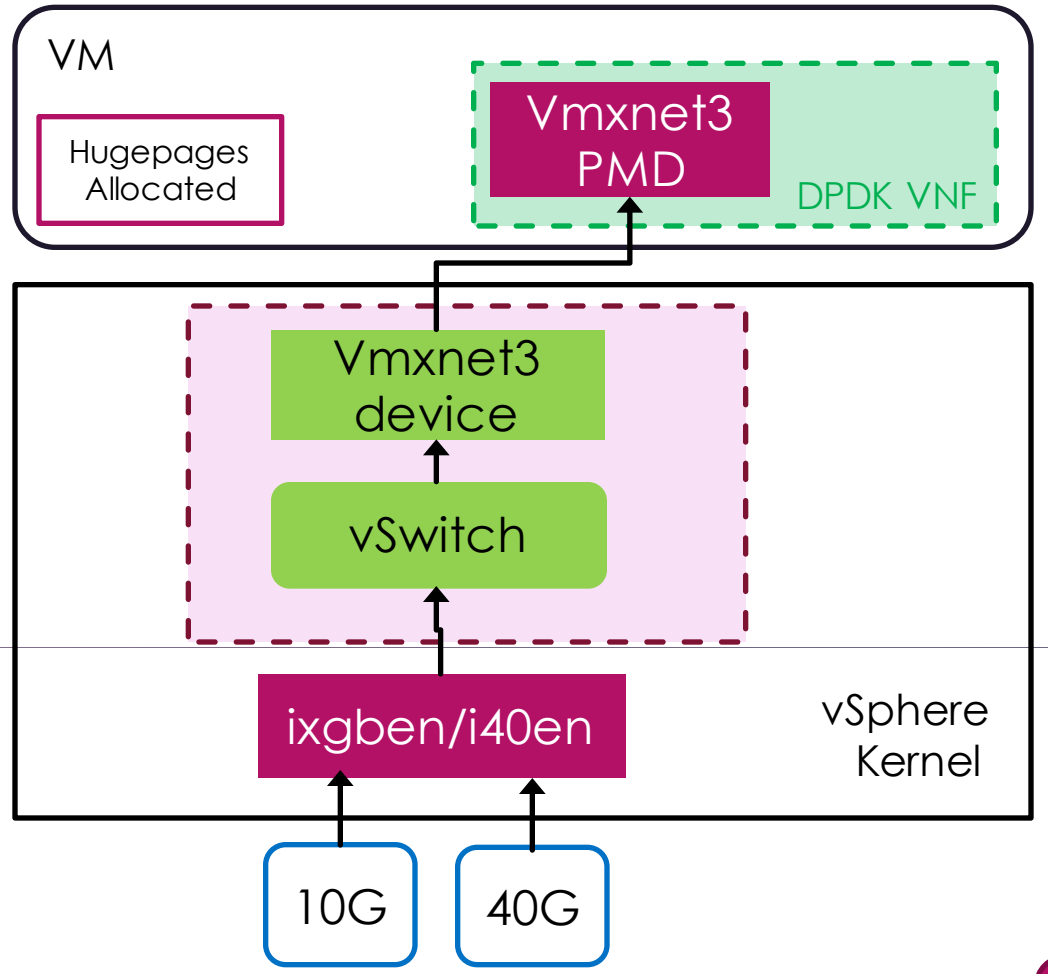
Accelerating NFV with VMware's Enhanced Networking Stack (ENS) and Intel's Poll Mode Drivers (PMD)

JIN HEO (HEOJ@VMWARE.COM)

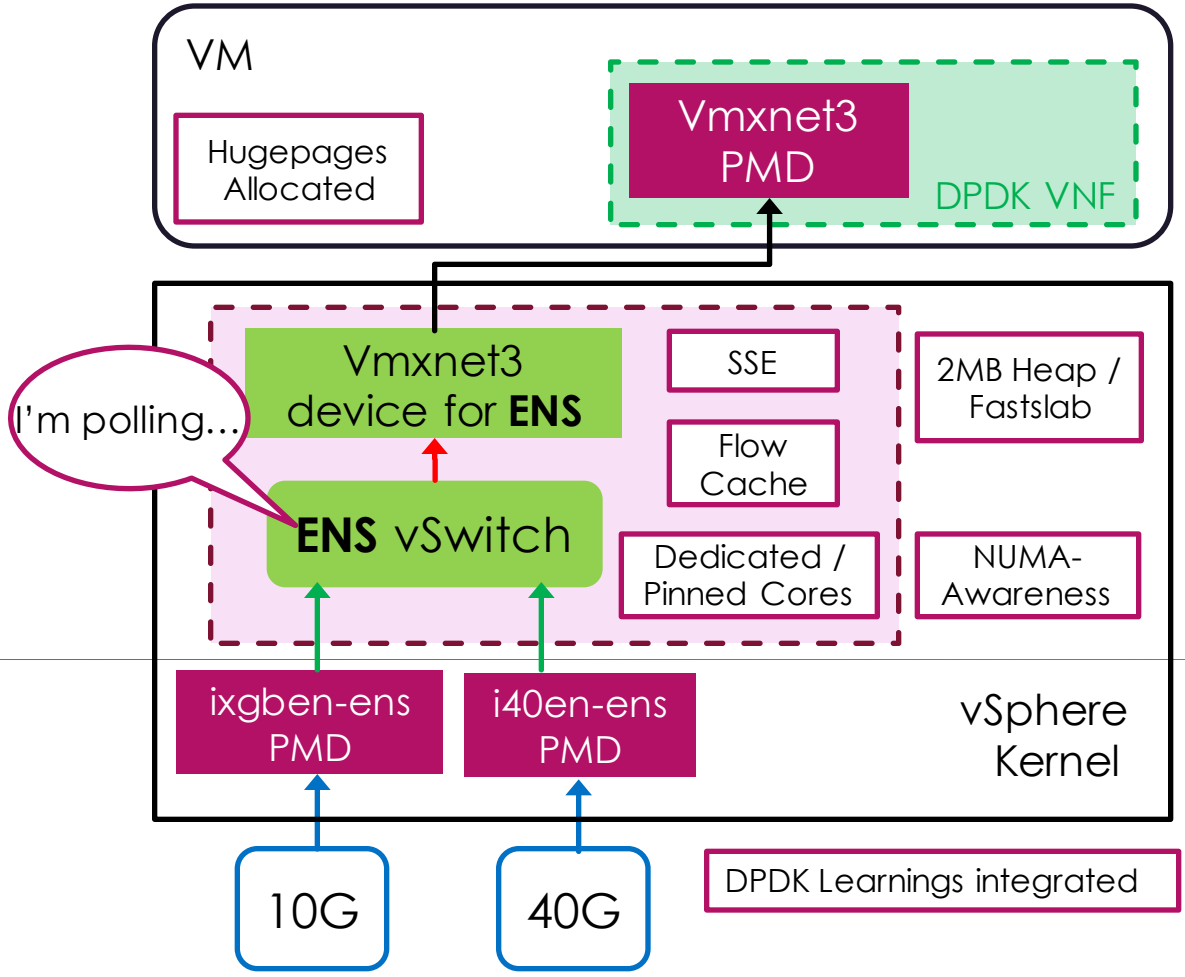
RAHUL SHAH (RAHUL.R.SHAH@INTEL.COM)

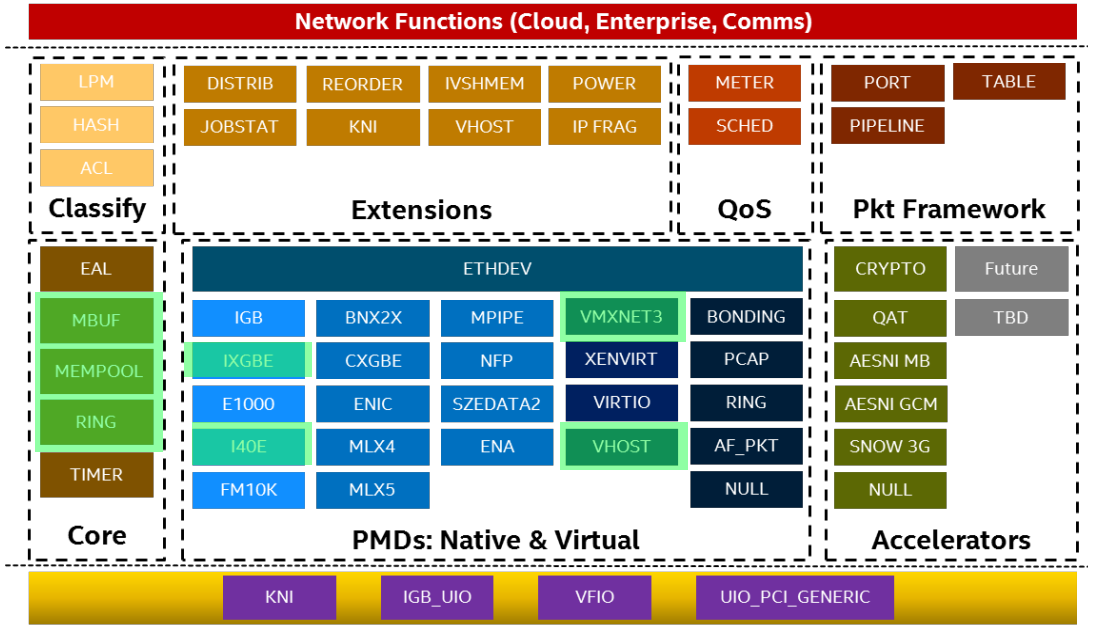
Problem Statement

- NFV has stringent performance requirements.
  - High packet rate for small packets
  - low packet loss and latency
- Need to improve vSphere networking stack to support NFV applications
  - Using DPDK in a VM is not enough.
  - Remove performance bottleneck in current networking stack
    - Vmxnet3 virtual device, virtual switching, and the physical driver
- We propose new **Enhanced Networking Stack** (ENS) for vSphere with **Intel poll mode physical driver**.



Solutions: Intel – VMWare Collaboration





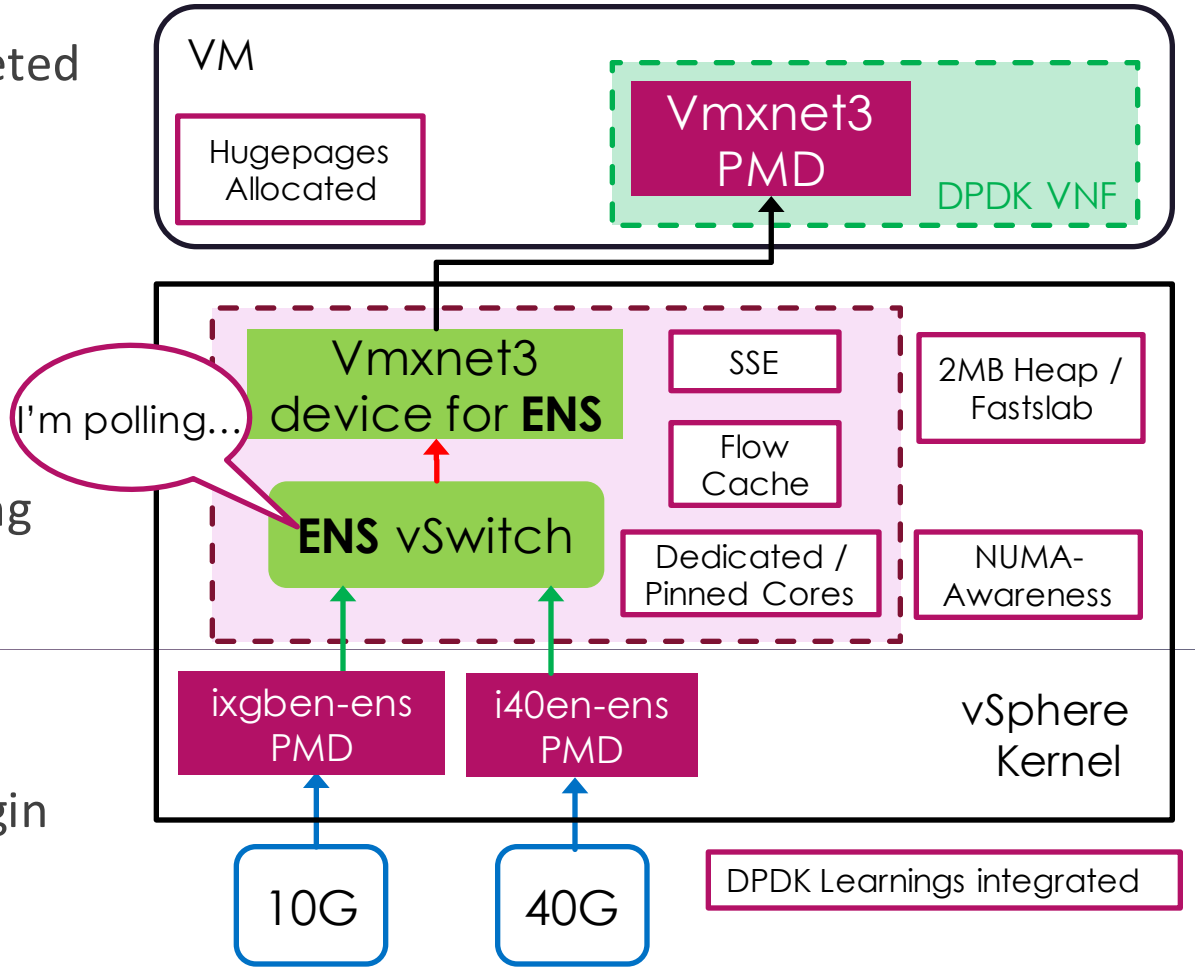
DPDK API Package

- Intel provided Poll Mode Drivers (ENS compliant)
- New ENS Driver Development Kit
- New ENS DataPath to VM

VMware Enhanced Networking Stack (ENS)

DPDK

- New and faster vSphere networking stack targeted for NFV applications
  - DPDK techniques employed
  - New vmxnet3 virtual device backend
  - New poll-mode physical device drivers
  - Faster switching using flow cache
- Deliver improved performance while supporting vSphere features
  - DRS, HA, vMotion
- Integrated with NSX
- Openstack (VIO) support through Neutron plugin



ENS Design Choices for Improved and Deterministic Performance

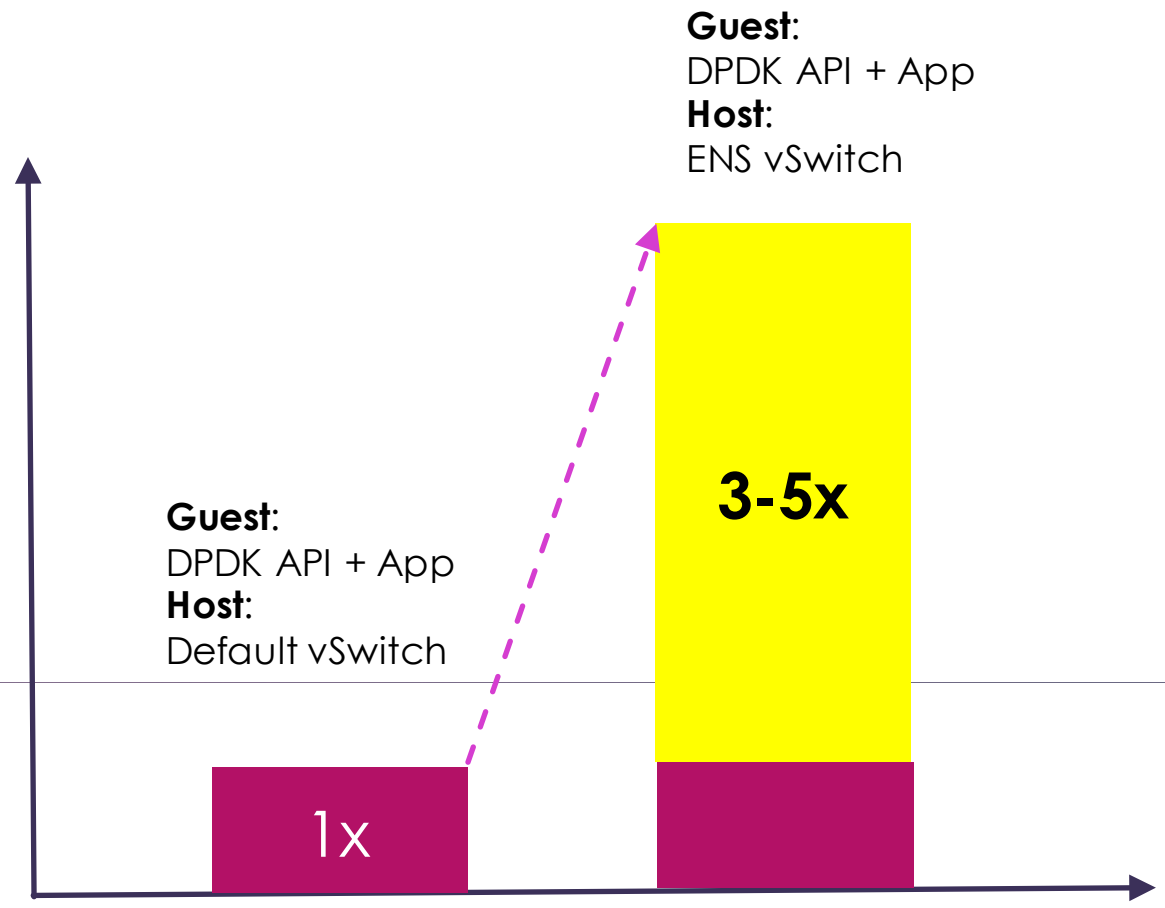
- Dedicated CPU allocation to system thread and polling
- NUMA-aware placement of VM and system threads
- NUMA-aware allocation with large pages
- Simplified packet representation
- Use of flow cache
- Lockless datapath
- Vmxnet3 optimizations
- SSE instructions faster packet processing

Intel ixgben-ens & i40en-ens PMD Release

- Initial ENS Poll Mode Drivers from Intel
  - IXGBEN-ENS
  - I40EN-ENS
- Initial Features
  - Receive/Transmit routines
  - Link Set/Get
  - Per Queue statistics
  - IPv4 TCP/UDP Checksum
  - Multiqueue filtering
  - Device reset

ENS Performance

- **3-5**x improvement in packet rate over the existing vSphere networking stack
  - Performance scales with the number of system threads
- Acceptable packet loss
- Low jitter and latency



“

Thank You

”