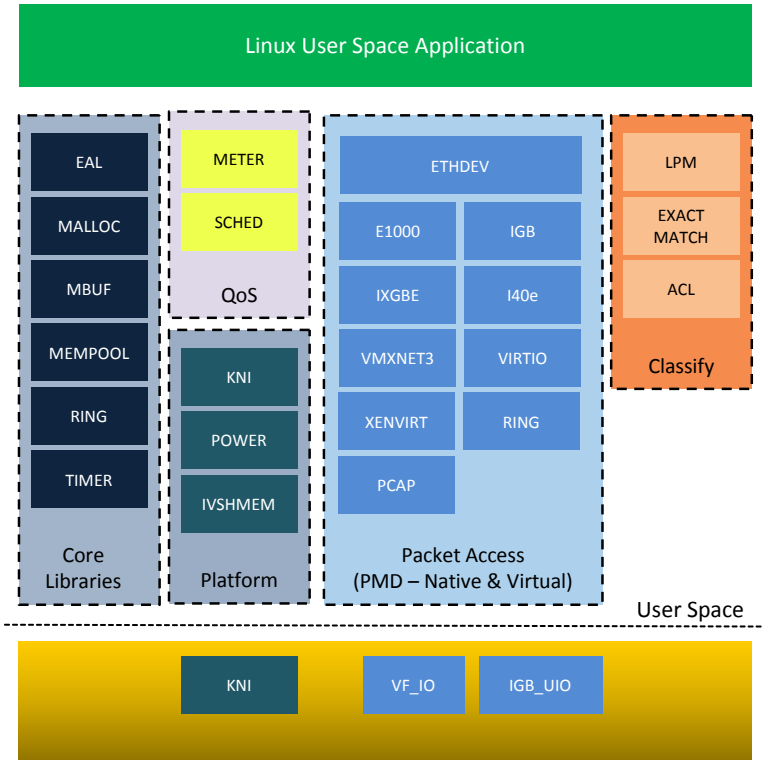
FUTURE ENHANCEMENTS TO DPDK FRAMEWORK

Keith Wiles, Principal Engineer, Intel Corporation

THE NEW CENTER OF POSSIBILITY

PRESERVING APPLICATION INVESTMENT WITH DPDK



- **Open-source (BSD license) community project (5+ years, version 2.1 latest) -- http://dpdk.org/**
  - All code is Open Source including the device drivers or PMDs (Poll Mode Drivers)
  - Optimized Linux User Space Library focused on data plane implementation on general purpose processors
  - Has been Very stable project with ABI versioning for APIs
  - Multi-architecture: x86, IBM, Freescale, EZChip(Tilera) support
- **Encompasses legacy platforms and newer acceleration platforms**
- **DPDK has a large application install base and included in Linux Distro's CentOS, Ubuntu, Red Hat, ...(Fedora)**
  - Adopted by standard OS distributions (FreeBSD, Linux) and many platform frameworks including VirtIO/Vhost and OpenvSwitch
- **Scalable solution to meet different NFV use cases**
- **Hardware acceleration complemented by software implementations for consistent set of services to applications**
- **Supports a large number of features like lockless rings, hash keys, ACL, Crypto, Match Action, buffer management and many others**
- **Has a large number of example applications and growing**

DPDK

DATA PLANE DEVELOPMENT KIT

DPDK - AE

What is Acceleration Enhancements for DPDK?

DPDK – WHAT DOES THE FUTURE HOLD?

Here are a few items we are thinking about and need help

- DPDK-AE (Acceleration Enhancements)
- What type of acceleration device types?
  - Crypto via hardware and software acceleration
  - DPI engine
  - Compression
  - Match Action and Flow Director APIs
- Adding support for SoC hardware
  - hardware memory management and event handling
- Network Stacks, light weight threading and other applications
- Focus on VirtIO performance and enhancements
- Support other language bindings

DPDK

DATA PLANE DEVELOPMENT KIT

DPDK – CRYPTO API

Overview of proposed Crypto API for DPDK

DPDK – CRYPTO USING HARDWARE AND SOFTWARE

Doing hardware and/or software crypto has some good advantages

- Hardware crypto can handle the large packets
- Software crypto can handle the smaller packets

Added advantages are:

- better performance over a range of packet sizes
- parallel execution with software and hardware crypto
- Abstracts the packet handling making it transparent to the application

DPDK

DATA PLANE DEVELOPMENT KIT

DPDK – FLOW CLASSIFICATION

Proposed flow classification support in DPDK

DPDK – FLOW CLASSIFICATION WITH HARDWARE

DPDK uses Flow Director APIs to manage flows

Match-Action API is a superset of APIs for flow classification

- The code is open source at https://github.com/match-interface

Match-Action API has a much large set of APIs to handle more flow classification needs, which we need to expose in the future

The Match-Action API is used under the Flow Director API for backward compatibility with current applications, while extending the applications to new hardware or software designs

DPDK – FLOW CLASSIFICATION WITH HARDWARE

DPDK uses Flow Director APIs to manage flows

The flows are currently managed in NIC devices, but we can extend FDIR APIs to support other hardware devices using Match-Action

Later we can continue to extend FDIR API to allow for more complex configurations and hardware designs using the full set of APIs with Match-Action APIs

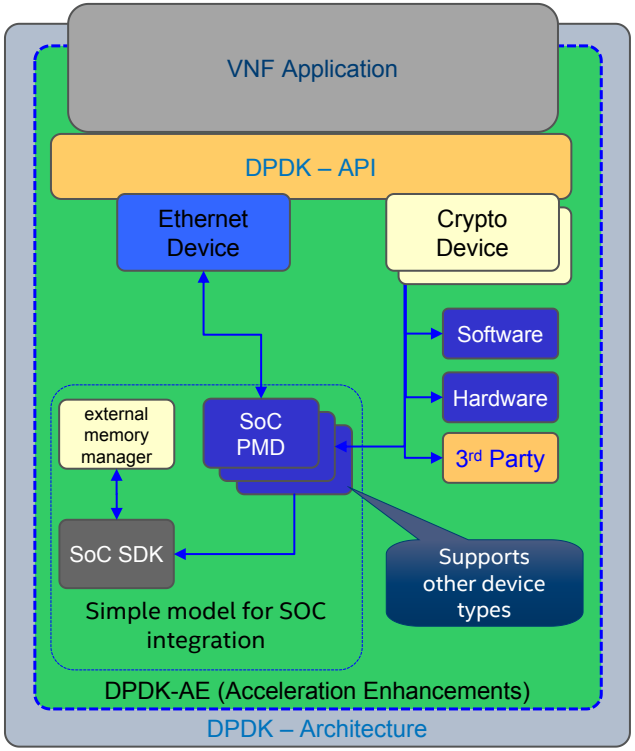
DPDK

DATA PLANE DEVELOPMENT KIT

DPDK – SOC SUPPORT

Proposed suggestion to add SoC support to DPDK

DPDK EXTENDING ACCELERATORS VIA SOC HARDWARE



SoC PMD: Poll Mode driver model for SoC devices

Provides a clean integration of SoC via a PMD in DPDK

- Hardware abstraction in DPDK is at the PMD layer
- DPDK-API: A generic API extended to support SoCs
  - DPDK provides a two layer device model to support many devices at the same time/binary, which can include SoC devices
  - Need to enhance DPDK with some SoC specific needs or features to support SoC hardware
    - Non-PCI configuration
    - External memory manager(s) (for hardware based memory)
    - Event based programming model
- SoC-PMD: Poll Mode Driver model for SoC

DPDK – CHANGES TO SUPPORT SOC HARDWARE

Enabling SoC hardware in DPDK requires a few enhancements

- Need a way to configure these non-PCIe devices
- Add support to DPDK mempool's to allow for external or hardware memory managers
- Add support for event based applications
  - e.g. Open Event Machine or others to utilize an event based programming model

Enlisting input for other enhancements to DPDK for SoC devices

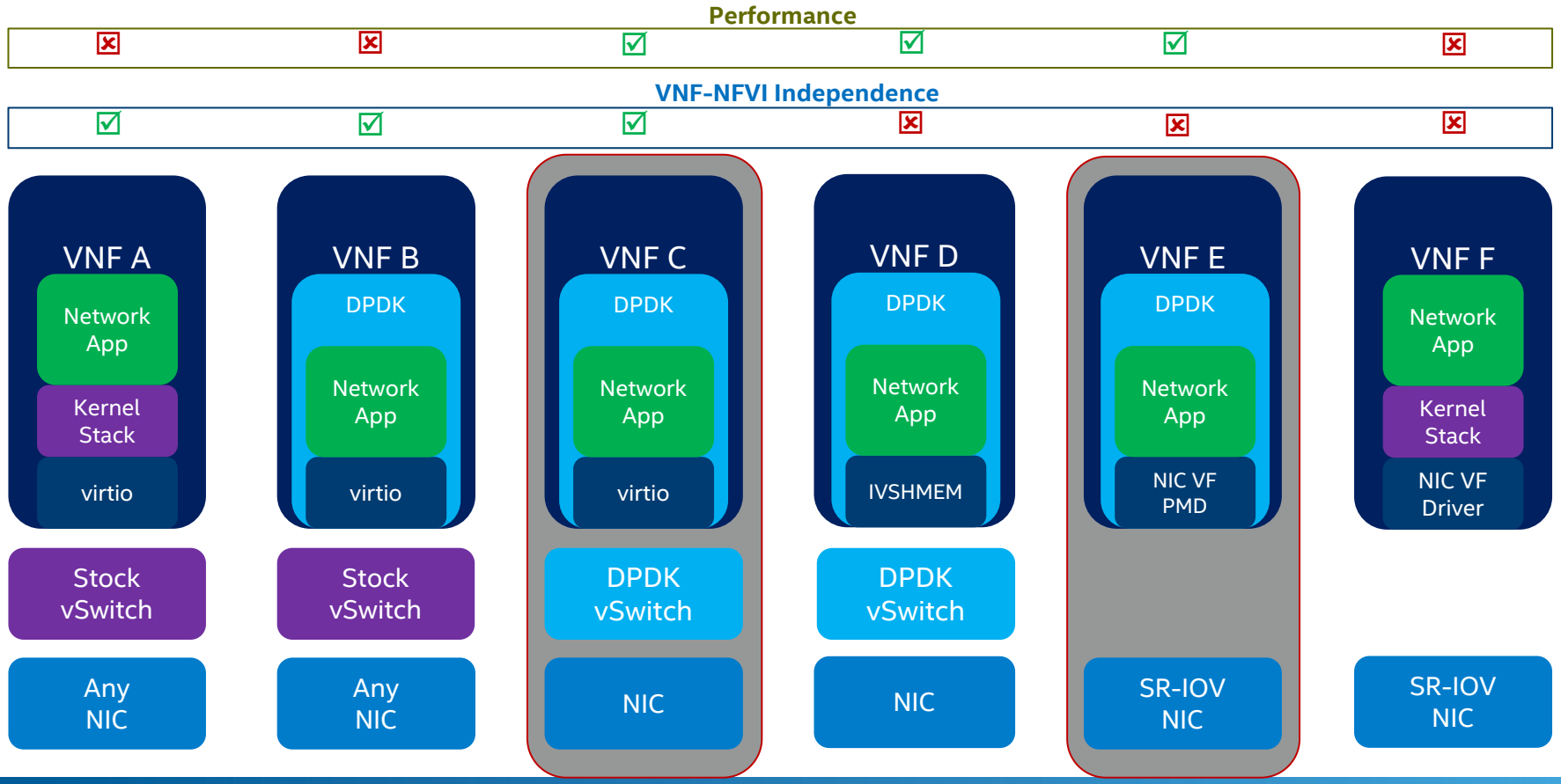
DPDK

DATA PLANE DEVELOPMENT KIT

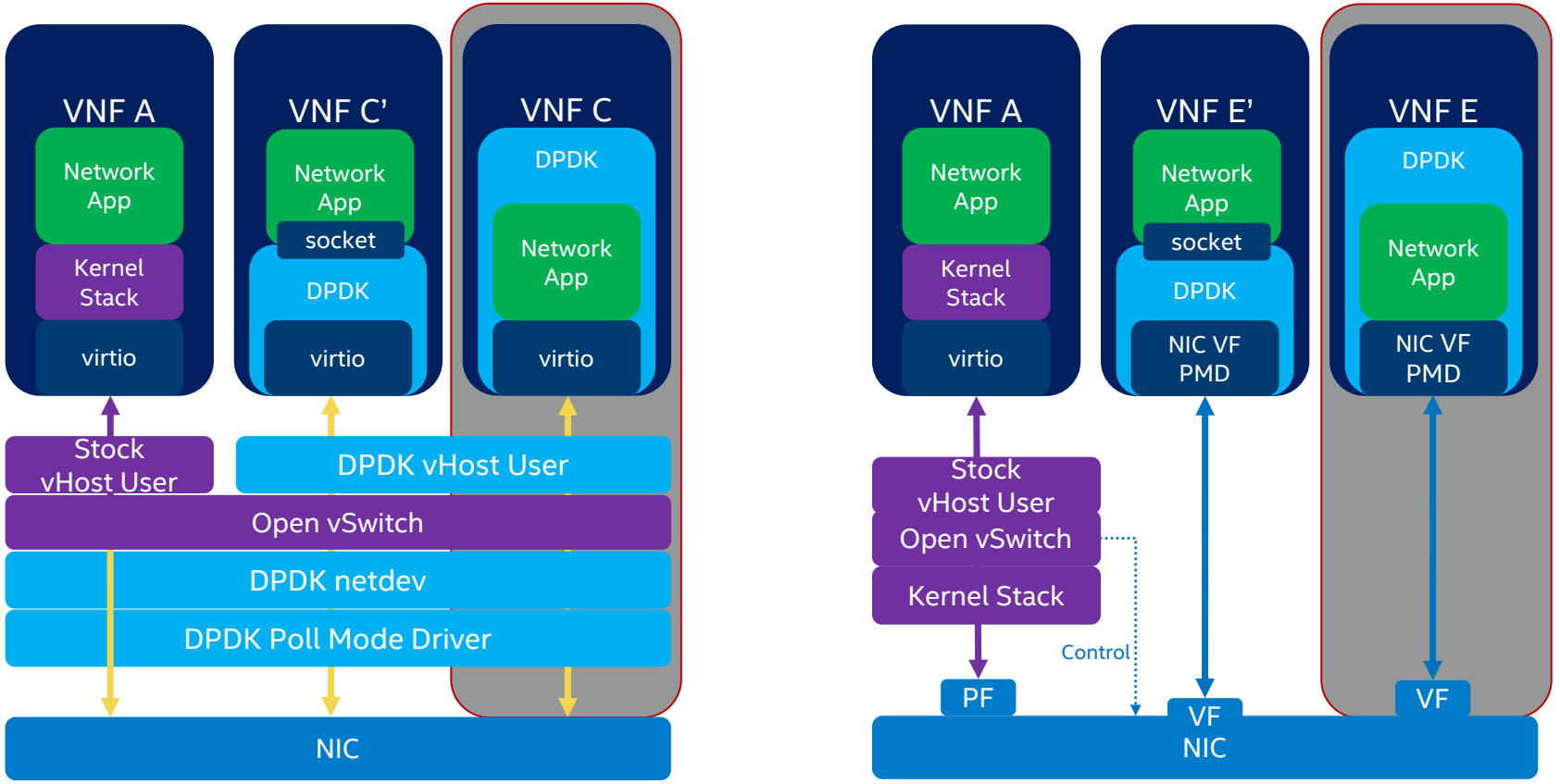
DPDK – NFV/VNF APPLICATIONS

Quick look at NFV/VNF applications in DPDK

VNF VIRTUAL NETWORK INTERFACE OPTIONS

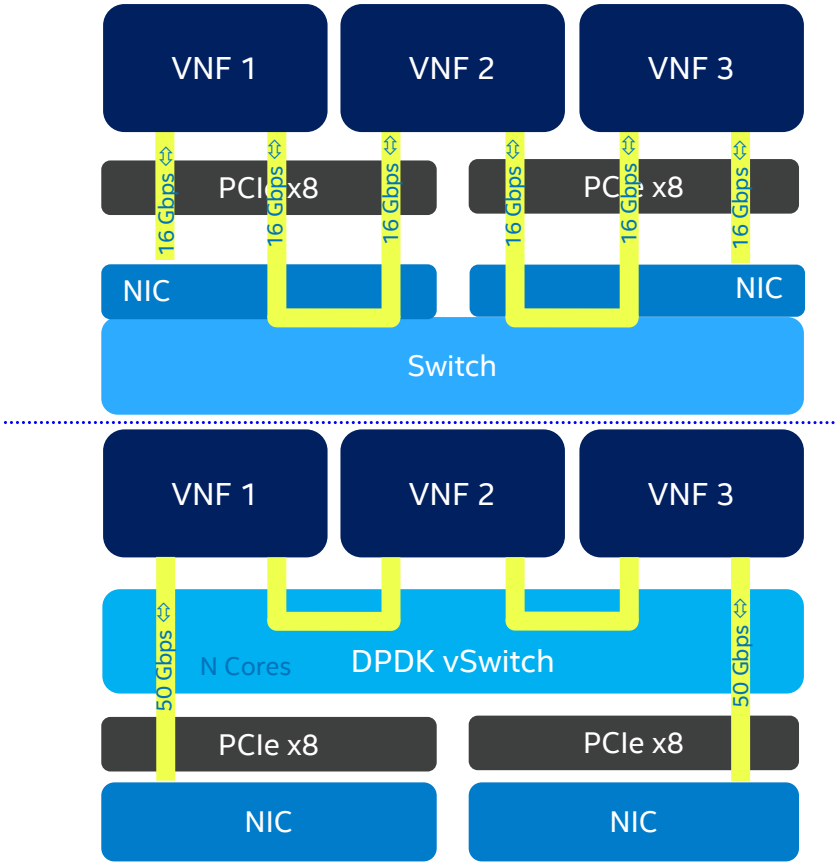


VIRTIO W/ DPDK OVS AND SR-IOV W/ OVS



SERVICE CHAINING BOTTLENECKS

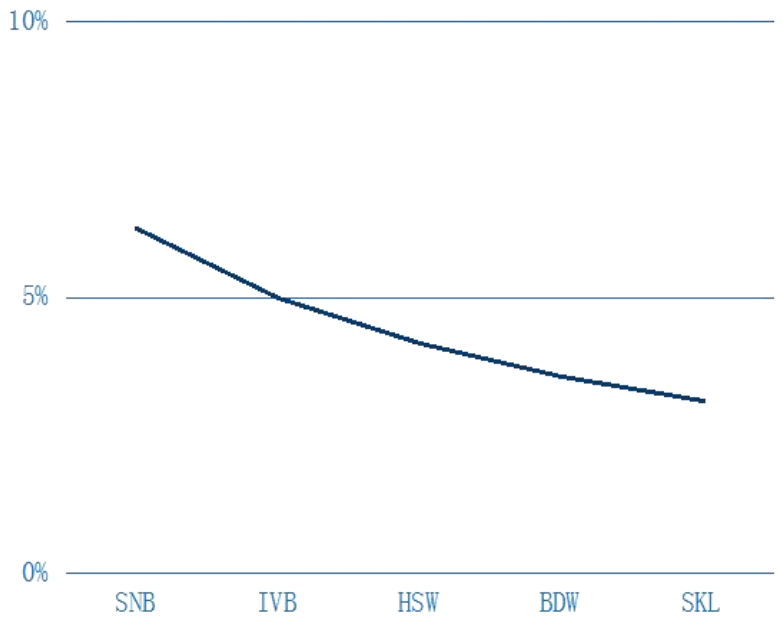
- Always need to consider the overall system performance, including internal bottlenecks
  - Having multiple VNFs on the same platform as part of a service chain leads to much greater VM-VM traffic than has been typically foreseen
- Supporting a 3 element service chain through PCIe x8 Gen 3 would limit throughput to 16 Gbps (50 Gbps/3)
  - Can scale number of cores to meet VM to VM traffic needs, (number needed will depend on packet size)



PAYING THE "CORE" TAX OF A VSWITCH

- We have seen (and will continue to see) core counts increasing over time
- As long as the core count remains constant, or increases slowly, the additional burden of dedicating a small number of cores becomes less of an issue
- e.g. using 2 cores to achieve high performance switching

$$\lim_{x \to \infty} \frac{2}{x} = 0$$



What is the Preferred NFV Solution?

External vSwitch

Pros
- Performance limited only by silicon
- Lowest latency/jitter

Cons
- No support for Live Migration
- Ties VMs to HW
  - issue of longevity and placement
- Difficult to extend capabilities
- Fixed TCAM size
- Easy to replace our solution with competitor

vSwitch Acceleration

Pros
- Leverages system architecture to its fullest
- Can extend DPDK-AE for other capabilities (e.g. crypto)
- Virtual interface supported in all VMs now
- Live Migration supported
- Longevity of VMs into the future
- Can be extended to support containers
- Can inspect/modify packets by hypervisor, add new features
- Scalable (#VMs, #Flows)
- Can adapt to different NIC capabilities

Cons
- Challenges in meeting line rate at 100G
- Latency and jitter needs to be optimized

vSwitch Acceleration is the most optimal solution for a scalable NFVi

DPDK – SUMMARY

- We need to add more acceleration supported hardware
  - Review and comment on the Crypto RFC
- Adding SoC enhancements to DPDK for more devices
- Adding better support for VNF/NFV applications is needed
- Creating a complete top to bottom NFV solution is the goal

- Lets collaborate on these and more…

DPDK

DATA PLANE DEVELOPMENT KIT

BUILDING A COMMON PLATFORM FOR EVERYTHING AND EVERYWHERE!

THANK YOU