USER SPACE NETWORKING USING DPDK AND ODP

RAVI MALHOTRA
PRODUCT MARKETING

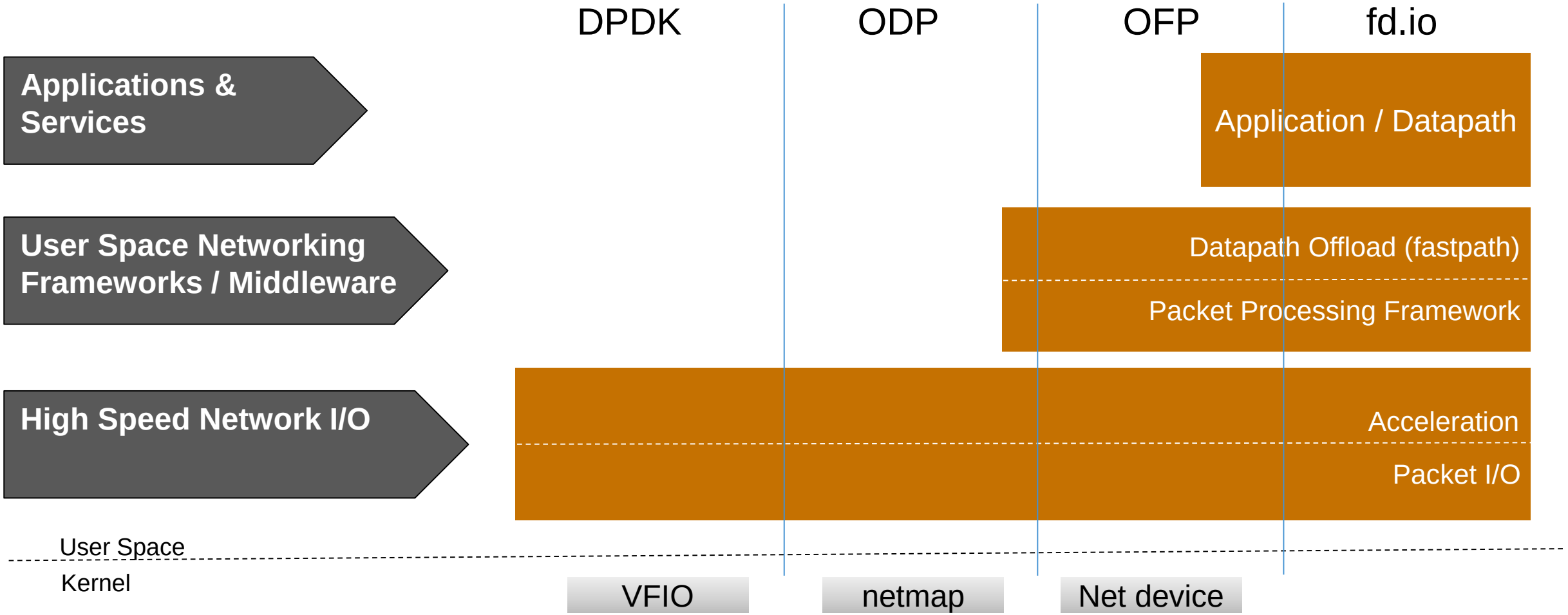
AMF-NET-T2674 | JUNE 2017

SECURE CONNECTIONS FOR A SMARTER WORLD

AGENDA

- Need for a common data-path API
- DPDK and ODP
  - What's special about them
- Choosing between DPDK and ODP

Key Open Initiatives for User Space Networking



User space network allows network I/O and packet processing frameworks to co-reside with Application, resulting in improved performance, flexibility and agility

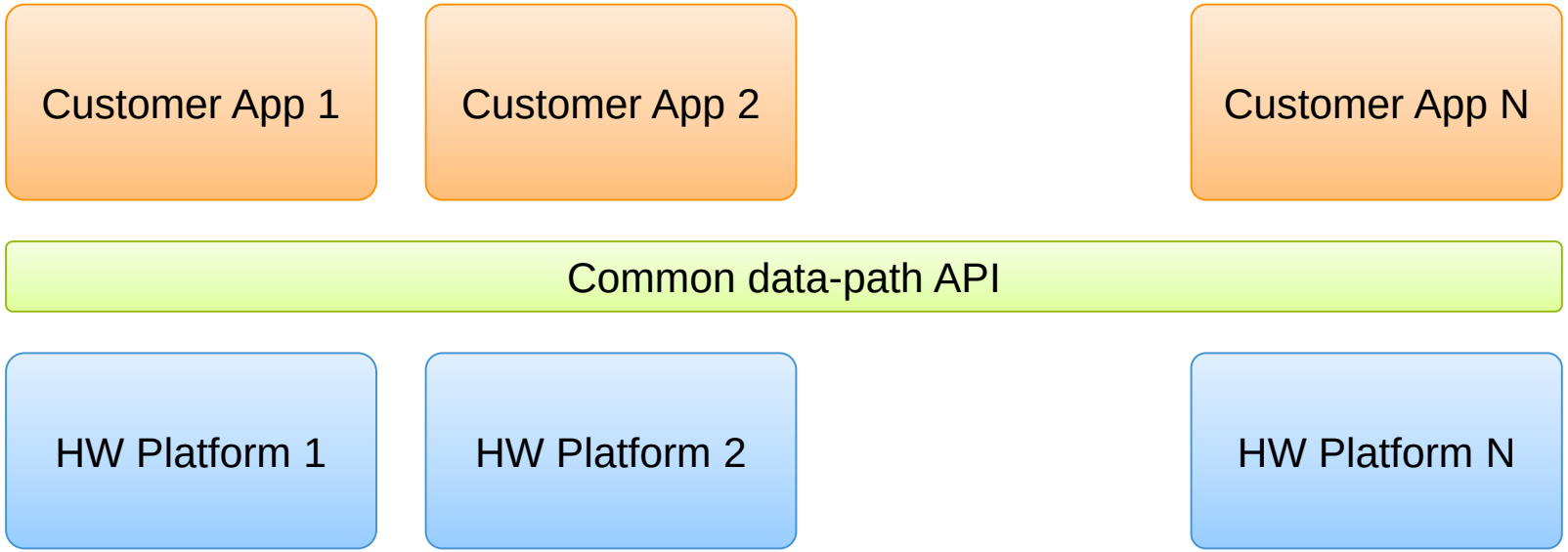
Traditional User-space Offerings – Vendor Proprietary

Vendor	Offering	Platform	Year introduced
Broadcom	Hyper-Exec, NetOS	XLR, XLP, XLS	2004
Cavium	Simple-exec, US App layer	Octeon	2005, 2009
NXP (Freescale)	Lightweight-exec, USDPAA	QorIQ DPAA	2008, 2009
LSI	Run-time environment	Axxia	2010
Intel	DPDK	x86	2011

- **Traditionally, user-space offerings evolved from bare-metal counterparts**
  - Very low-level API
  - Designed for highest performance, and not for ease-of-use or portability

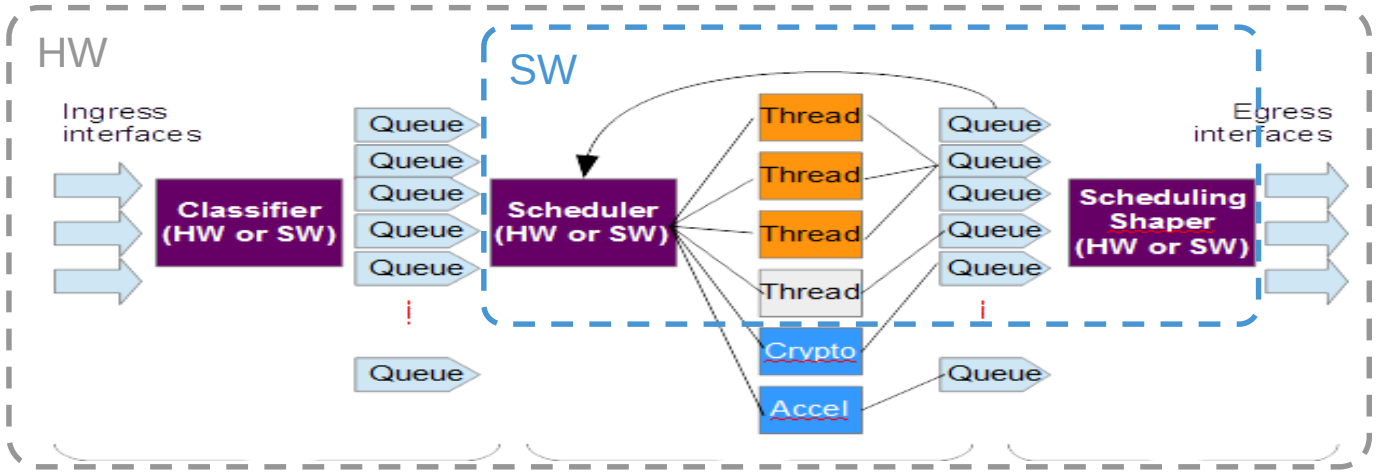
- **Use-cases were targeted – e.g.**
  - Routing/Gateway fast-path
  - Base-band transport and L2/L3 processing

Need For a Common Data-path API



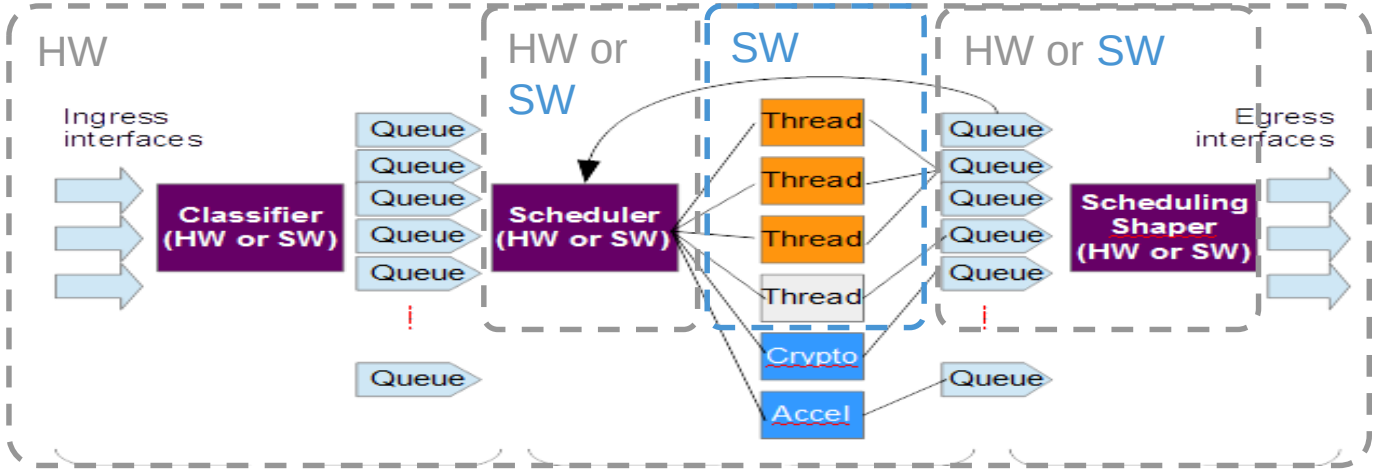
- **A common API**
  - Increases portability of applications across several HW platforms
  - Increases the number of applications that can run on a HW platform.
- **Is it possible, even probable?**
  - Basic I/O, acceleration and run-time services – Yes.
  - HW vendors will continue to add differentiation, value-added services for advanced functionality.
  - Provisioning and management also needs standardization – especially for NFV deployment.

DPDK vs. ODP – HW Acceleration



DPDK Approach:

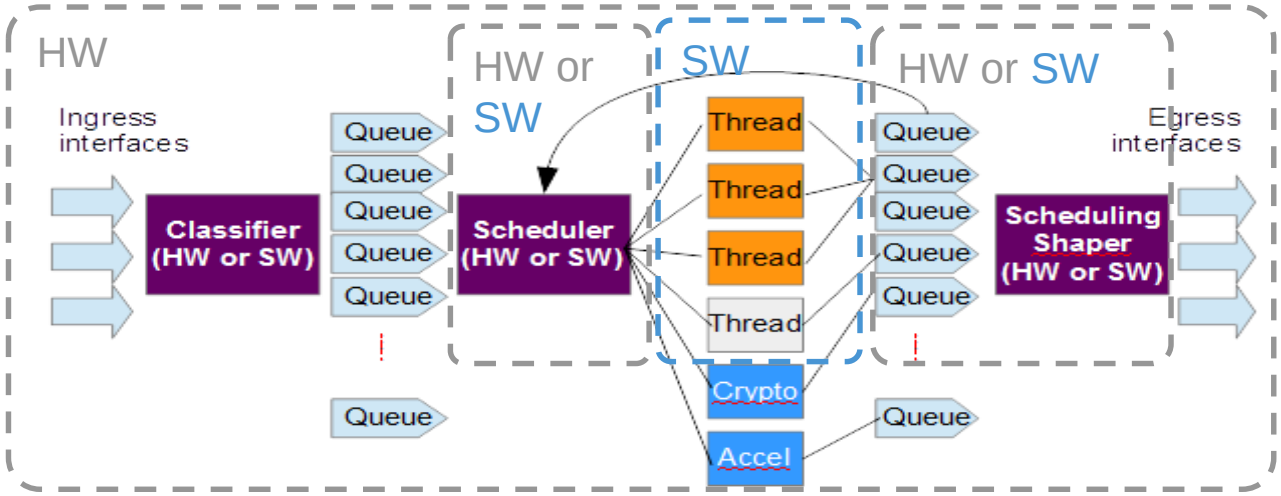
- Designed for Simple NICs
- Works well for large number of balanced flows
- SW implementation comes at a cost.

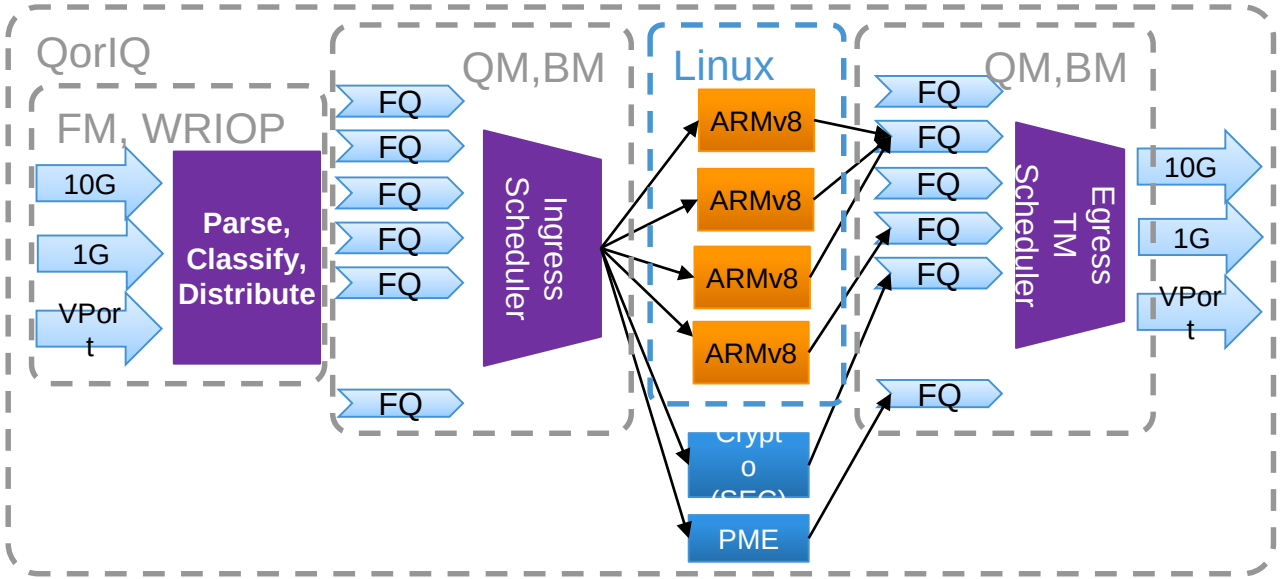


ODP Approach:

- Flexible design – blocks can be in HW or SW
- Works for balanced & unbalanced traffic flows
- Works well with Accelerators, multiple I/O sources

DPAA – Compatible With DPDK & ODP Since 2008





DPDK/ODP Approach:

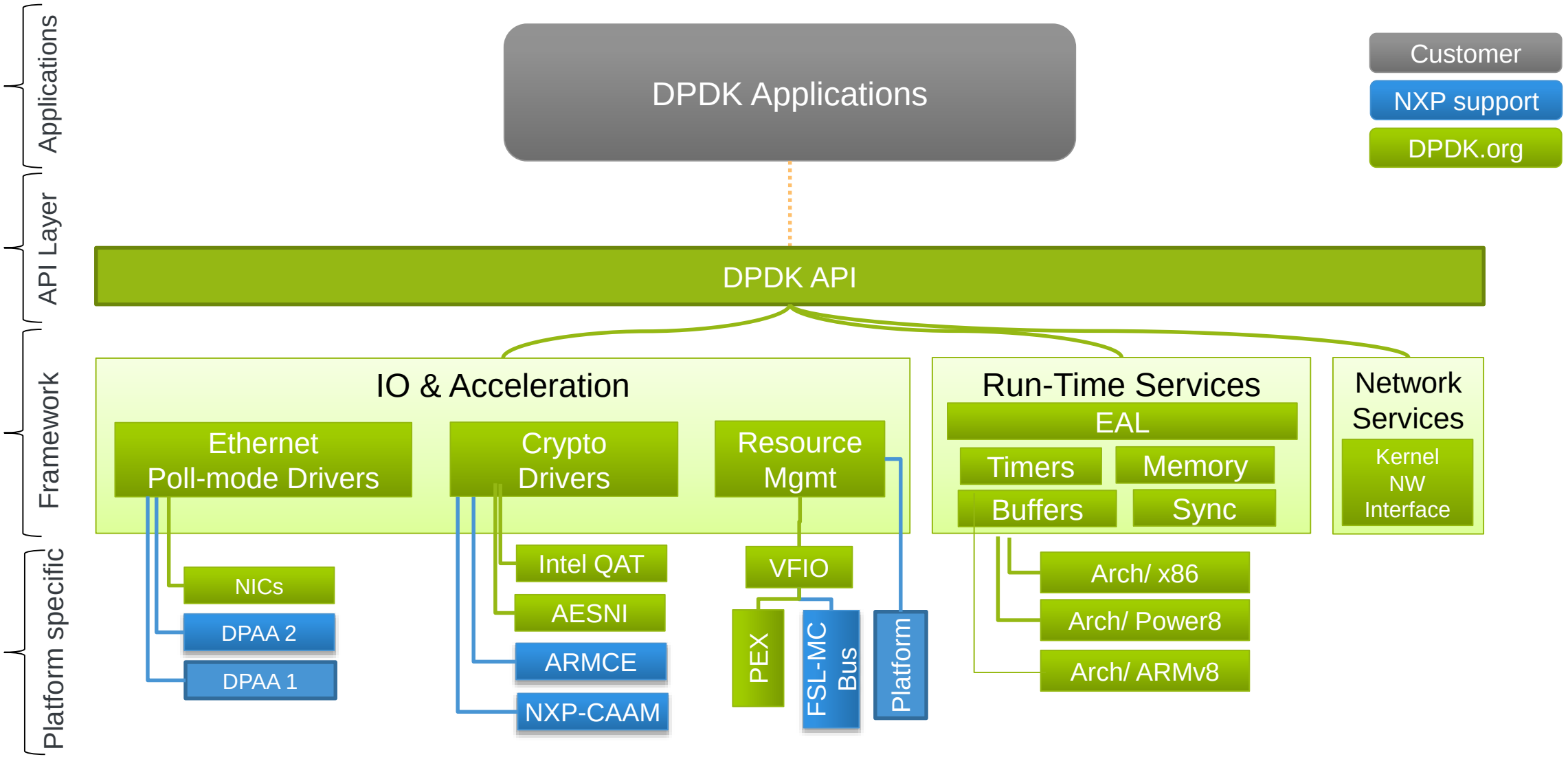
- Leverage hardware for ingress and egress processing
- ODP adds on HW scheduling offload
- Accelerators – Crypto offload
- Complete user-space processing model

NXP Approach:

- FMan offloads parsing, classification, distribution.
- QMan, BMan offload scheduling, buffering
- Virtualized accelerators – SEC, PME, DCE
- User-space driver, threading model
- Doing all this since 2008 – now into 3rd generation

01.
DPDK

NXP - QorIQ DPDK Support



* Efforts are underway to make NXP support part of standard dpdk.org

NXP - DPDK Support

- **Basic Platform support**
  - DPAA1 Poll-mode driver
  - DPAA2 Poll-mode driver
  - Crypto offload to SEC HW
  - Crypto offload to ARM NEON Crypto Instruction driver.
  - Platform supported
    - DPAA1 - S1043, LS1046
    - DPAA2 - LS208x, LS1088

- **Ethernet features**
  - Queue Management
  - RSS Hash, packet type parsing
  - Checksum offload,
  - MAC filter, Multicast/Unicast
  - Jumbo and SG support
  - Link Status event.

- **Architectural enhancements**
  - Support for SoC/platform drivers
  - Hardware buffer management
  - Optimizations for ARMv8 run-time

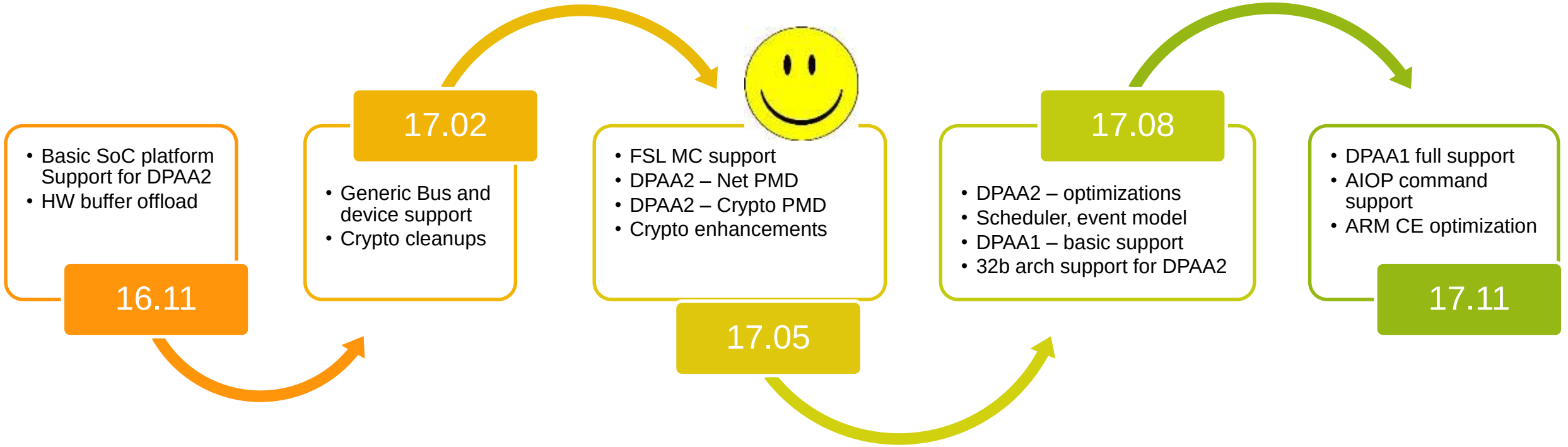
- **Virtualization support**
  - OVS over DPDK in host user-space
  - DPDK in guest/VM
  - Virtio poll-mode driver
  - DPAA2 VFIO poll-mode driver*

- **Application support**
  - L2fwd, L3fwd, testpmd
  - IP_fragmentation/IP_reassembly
  - l2fwd-crypto, ipsecgateway,
  - KNI , PKTGEN
  - FD.IO

DPDK Crypto Subsystem

- Session-less Mode
  - For each job, software defines;
    - The data to be operated upon (input buffers, lengths, offsets)
    - The output buffers to hold results
    - The cryptographic operations to be performed
    - Keys & context for the cryptographic operations
- Session Oriented Mode
  - For each job, software defines;
    - The data to be operated upon (input buffers, lengths, offsets)
    - The output buffers to hold results
  - Cryptographic operations, keys & context are defined at session establishment time, and referenced for each job
  - **Current support is only for Synchronous Mode** in most example applications.
  - NXP is working on proposal to support async crypto session APIS.
- Supports virtual and physical crypto devices
  - Virtual Device (Software Implementation)
    - Intel AES-NI/vector operations
    - ***ARM NEON instructions ****
  - Physical Device (Hardware Accelerated)
    - QAT
    - ***DPAA-CAAM****
    - ***DPAA2-CAAM****
- Test Applications
  - L2fwd with crypto
  - ipsec forward application
  - Testpmd

DPDK Upstreaming – NXP Roadmap

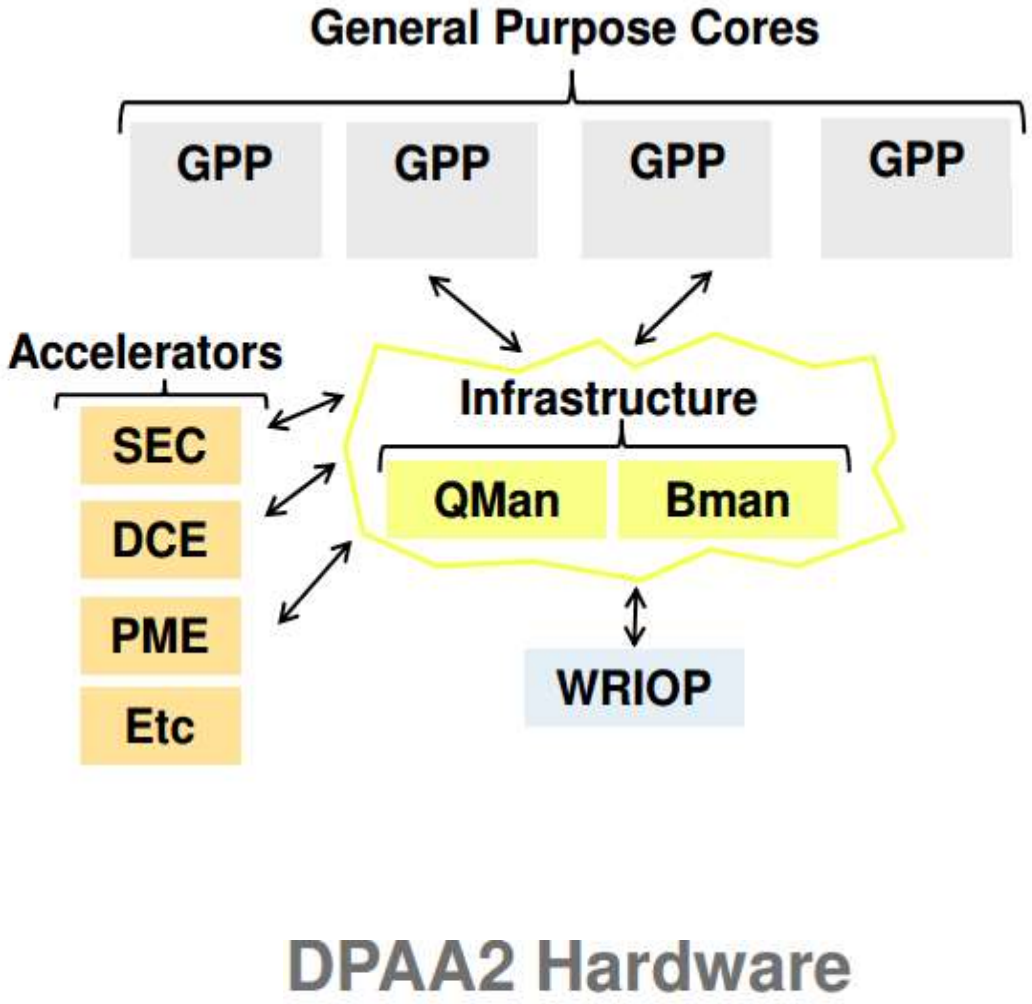


- ❖ Basic Support for 64 bit DPAA2 is now available in 17.05
- ❖ Performance patches are targeted for 17.08 release – under review.
- ❖ Changes in OVS to support hw offloaded pool are under review.

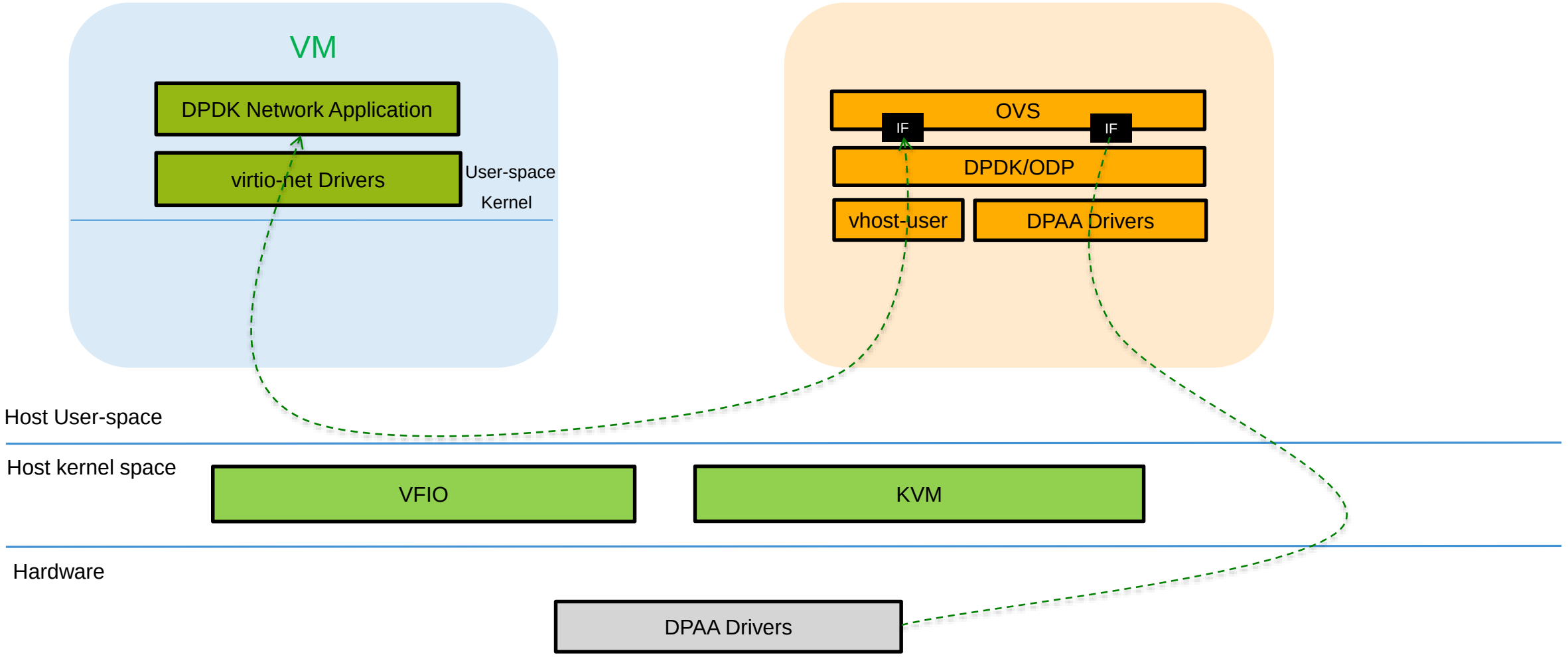
NXP is now a DPDK Technical Board Member (Rep: Hemant Agrawal)

DPDK on NXP SoCs

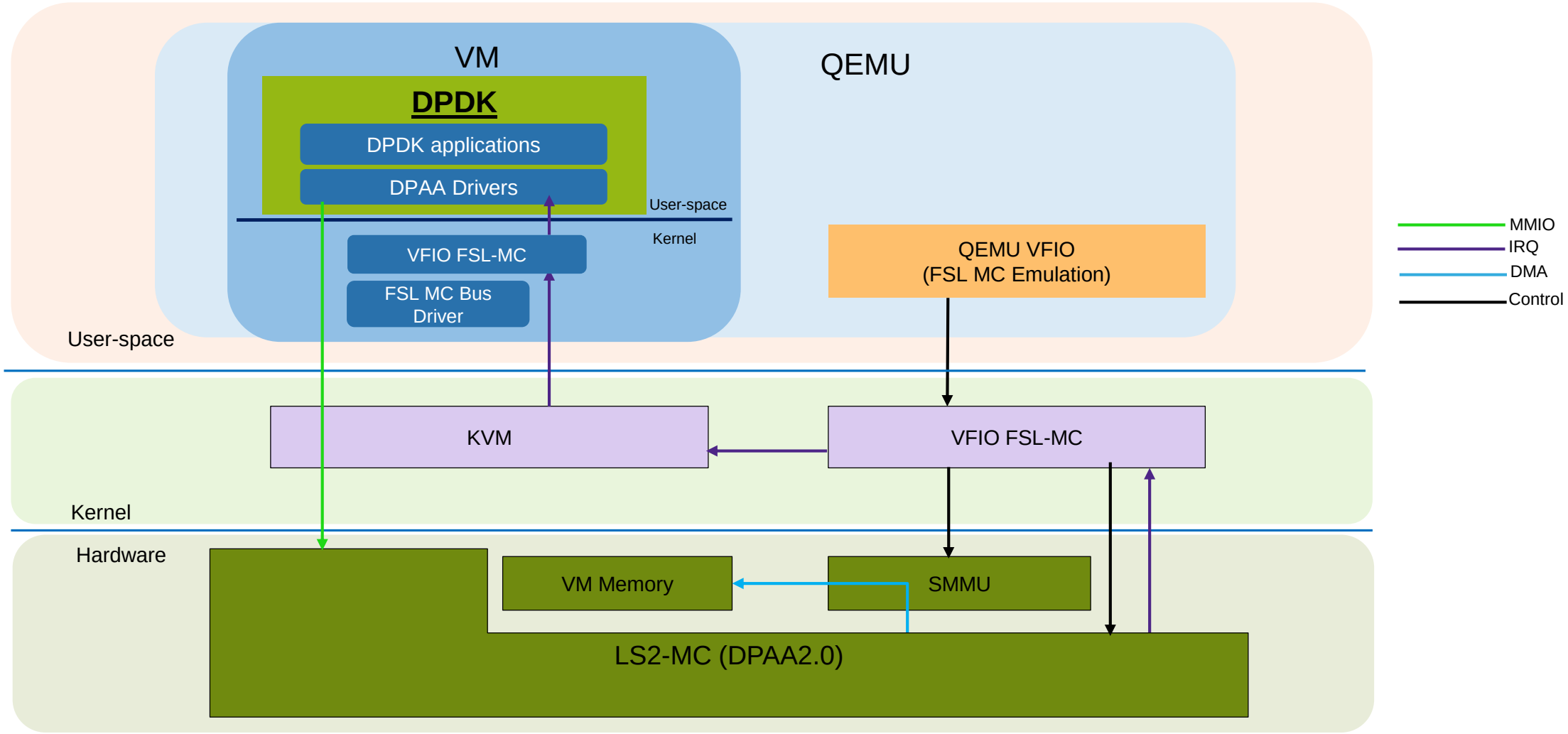
- DPDK 16.07 supports NXP platform configuration (without NXP PMDs)
  - `defconfig_arm64-dpaa2-linuxapp-gcc`
- NXP Networking SoC:
  - Have in-built MAC and they are non-PCI based
    - VFIO - FSL-MC BUS is used to scan devices.
  - Have in-built accelerators to support packet processing
    - BMAN - Packet buffer to be allocated & managed by HW
    - QMAN - Packet Queues mapped to hardware queues
    - CAAM – Crypto accelerator offload



Virtio-net: DPDK in guest using virtio-net



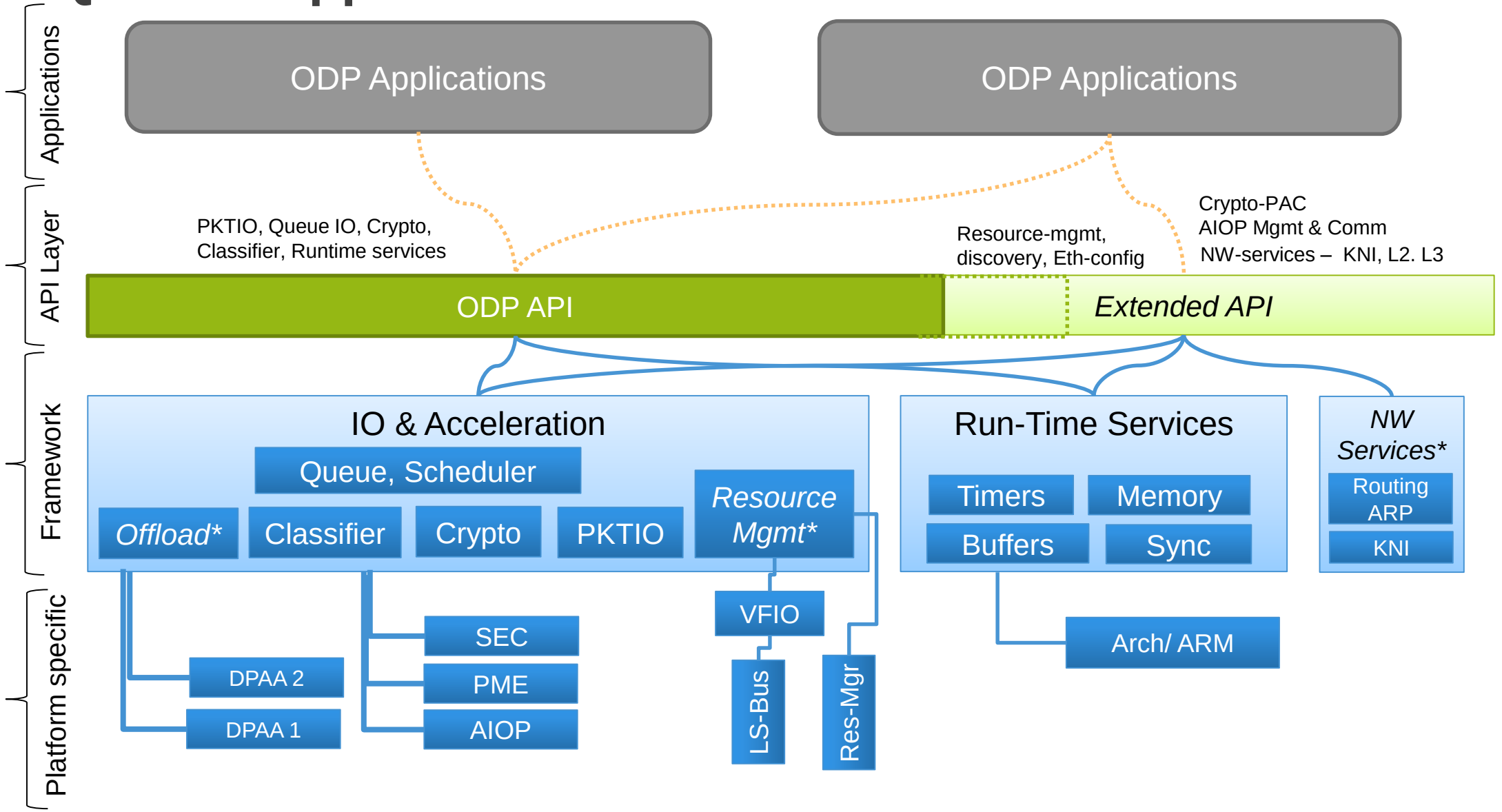
DPAA2 Device Pass-through to DPDK in VM



02.

ODP

QorIQ ODP Support

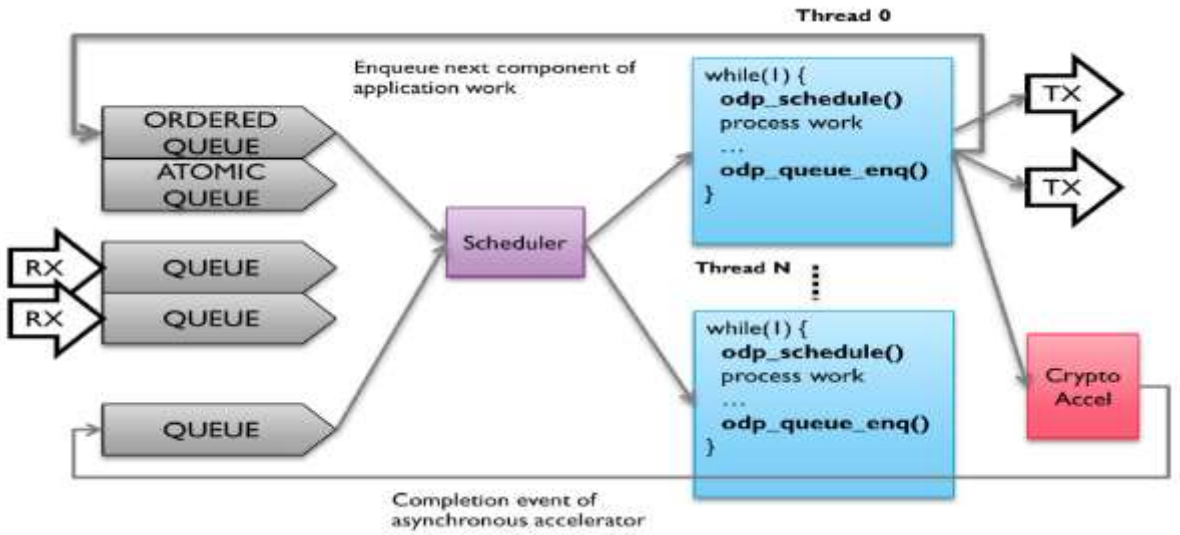


QorIQ ODP Support

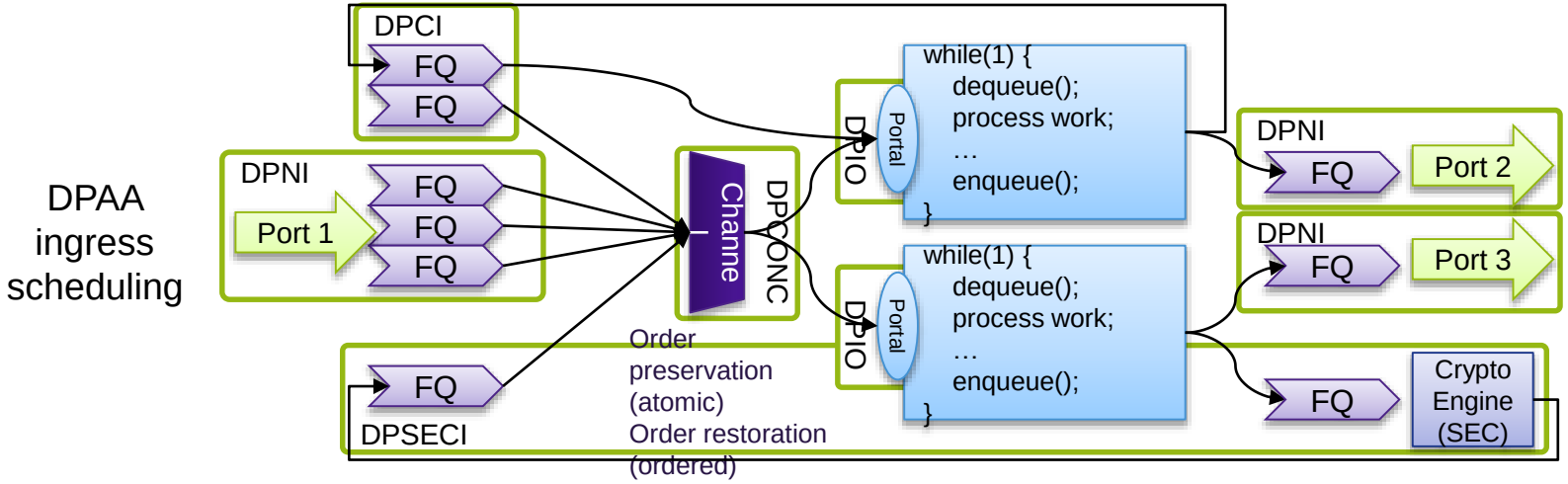
- **QorIQ HW is inherently aligned to ODP**
  - Classification and scheduling
  - HW queue and buffer mgmt
  - Crypto & other HW offloads
  - ARM 64-bit cores
- **Complete ODP-API coverage**
  - Queue and Scheduler API
  - PKTIO and Classifier API
  - Crypto API – algorithmic and protocol
  - Runtime services incl. pkt-buffers
  - Support for both DPAA1 & DPAA2 platforms
    - LS1043, LS1046
    - LS2088, LS1088
- **QorIQ HW have additional capabilities**
  - Switching, demuxing
  - Application level offloads
  - Virtual networking and resource mgmt
  - Provided as extensions to ODP-API
  - *Efforts underway to make them part of ODP*
- **Value-added ODP extensions**
  - Complete Ethernet capabilities
    - MAC/Phy, IPR/IPF, *GRO/GSO, Smart-NIC*
    - Physical and Virtual Ethernet ports
  - NW services
    - Provide Linux network stack services, visibility
    - Network-devices (KNI), Routing, ARP
  - Resource management
    - VFIO and VirtIO based assignment of resources.
    - Dynamic re-configuration and discovery
    - Multiple application support, flexible process model

Ingress Scheduling and Load-balancing in ODP

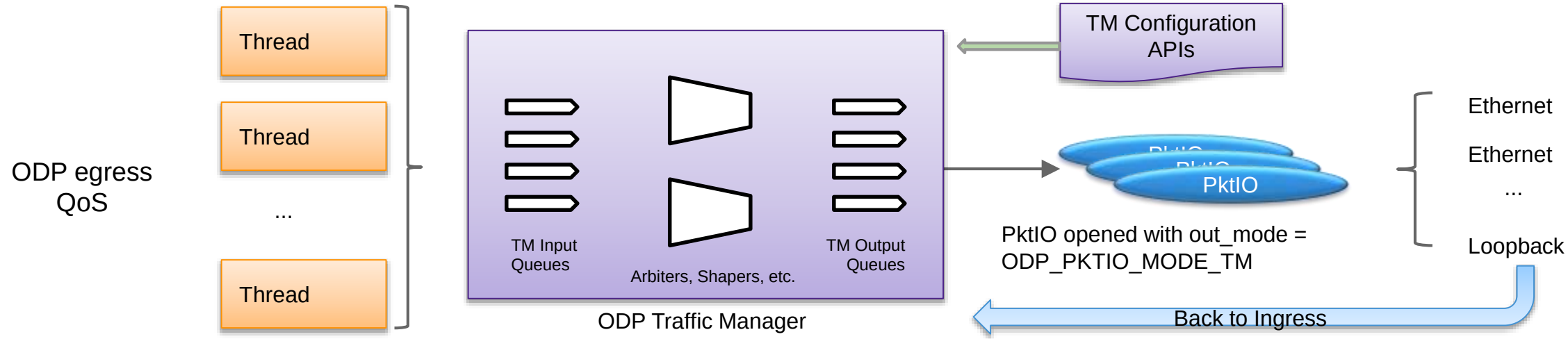
ODP ingress scheduling

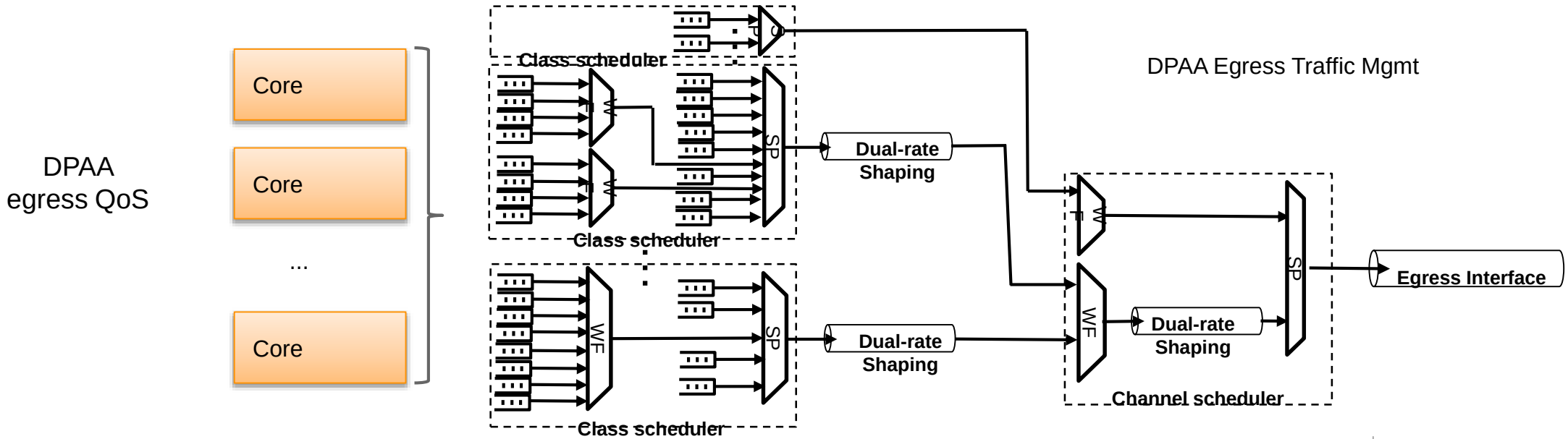


DPAA ingress scheduling



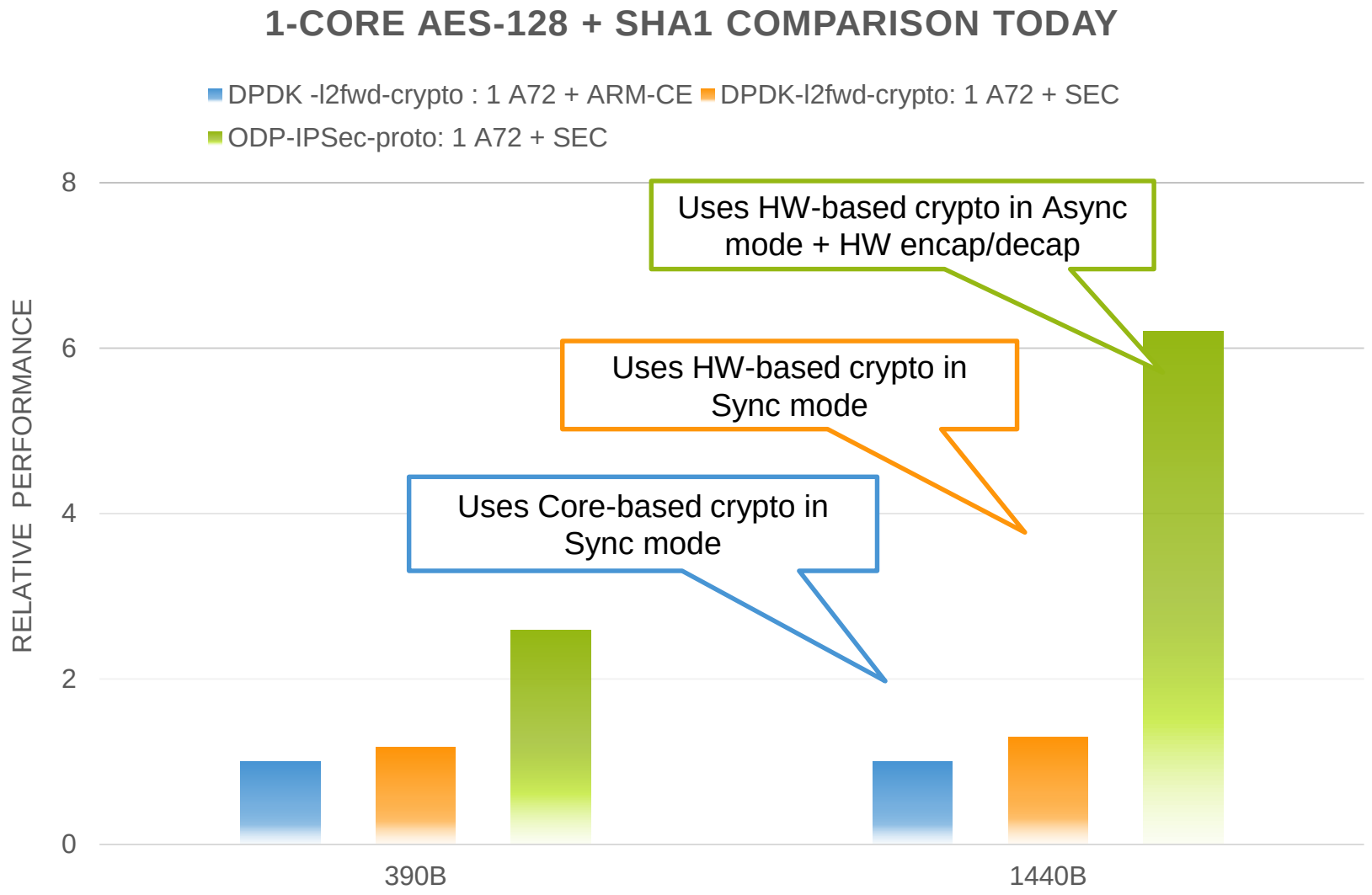
Egress QoS in ODP





ODP IPsec support

- ODP supports
  - Async HW-based crypto
  - IPsec encap/decap offload
  - Complete IPsec offload (planned)
- DPDK is catching up
  - NXP introducing scheduler support for async.
  - IPsec encap/decap also proposed.



03.

Choosing between ODP and DPDK

ODP or DPDK ?

	Data Plane Development Kit	Open Data Plane
Availability	Since 2012	Since 2014
Maturity	Wider ecosystem of applications	Limited ecosystem
Platform compatibility	Single dpdk.org release with support for all platforms.	Separate releases from each platform vendor
Virtualization	Rich support – VirtIO & Direct-assignment	In planning.
Load-balancing	Done in software	Supports in software and offload to hardware
QoS	Done in software	Supports in software and offload to hardware
Crypto	Both core-based (AES-NI, ARM-CE) and offload (SEC, QAT). Async mode, protocol offload – in planning.	Supports only offload (SEC) – both sync and async modes with protocol offload.
Inline and look-aside acceleration	No plans	In planning
NXP support	NXP supports both ODP and DPDK NXP is a key member of both ODP and DPDK and are driving the definition	

Summary

- **Need for a common user-space API**
  - Mainly driven by NFV and SDN
  - Best of portability, re-use and acceleration

- **Open Data Plane and Data Path Development Kit**
  - Different origins, communities - but lot of convergence
  - Both will continue to be adopted

- **NXP provides optimized solutions for both ODP and DPDK**
  - Our Data-Path architecture has been compatible since 2008
  - Working with the community to add more acceleration, features
  - Actively engaged in and tracking FD.IO and Open Fastpath communities

SECURE CONNECTIONS
FOR A SMARTER WORLD