Intel® Network Builders Insights Series

Introduction to a Cloud Native Approach for Packet Processing with Kubernetes CNI
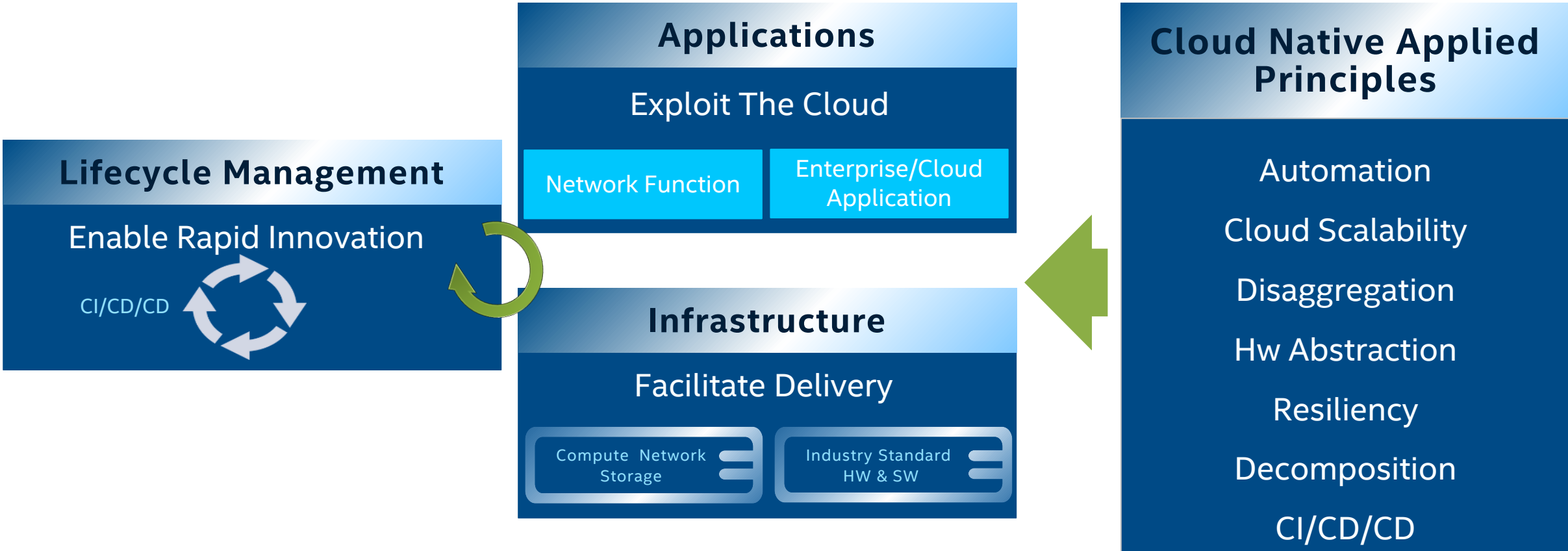
- Xiaojun (Shawn) Li, Sales Director, Next Wave OEM & eODM

- Brian Skerry, Principal Engineer

intel.

Notices and Disclaimers

- Intel technologies may require enabled hardware, software or service activation.

- No product or component can be absolutely secure.

- Your costs and results may vary.

- © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Cloud Native Delivery – What Does it Mean?

Applications

Exploit The Cloud

| Network Function | Enterprise/Cloud Application |

Cloud Native Applied Principles

Automation

Cloud Scalability

Disaggregation

Hw Abstraction

Resiliency

Decomposition

CI/CD/CD

Lifecycle Management

Enable Rapid Innovation

CI/CD/CD

Infrastructure

Facilitate Delivery

Compute Network Storage

Industry Standard HW & SW

Implement Principles Aligned With CNCF Cloud Native Definition Across Network Foundational Elements

The Evolution Toward Cloud Native

New Services Rapid **Introduction**

Operation Agility

Resources Efficiency

Improved End-user Experience
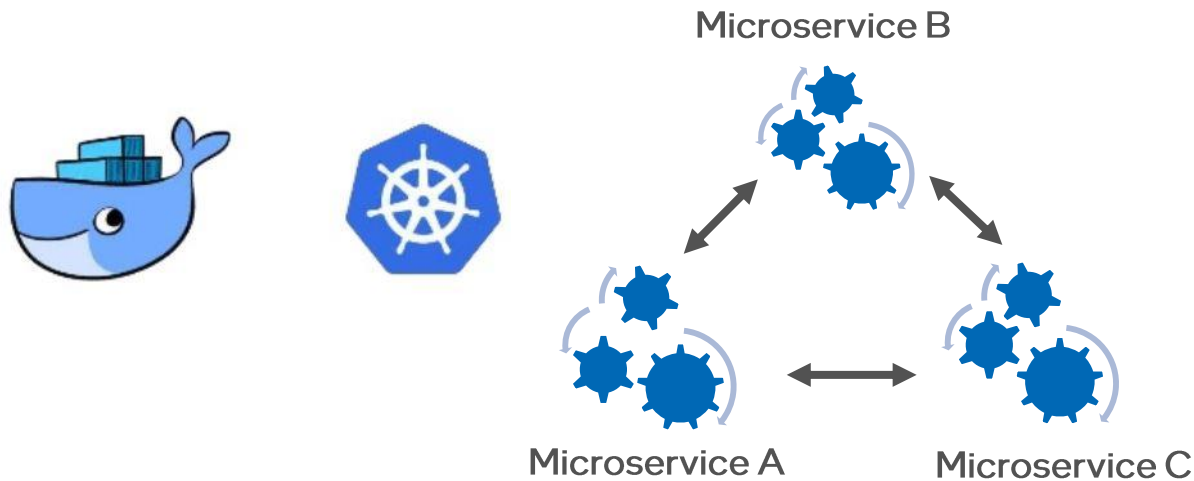
Comms Cloud Enterprise

5G

CLOUD NATIVE

Same Goals As With NFV – Builds on Years Of Industry Experience and Utilizes Cloud Native Technologies to Accelerate Achieving Industry Goals

Driving Rapid Innovation with Cloud Native
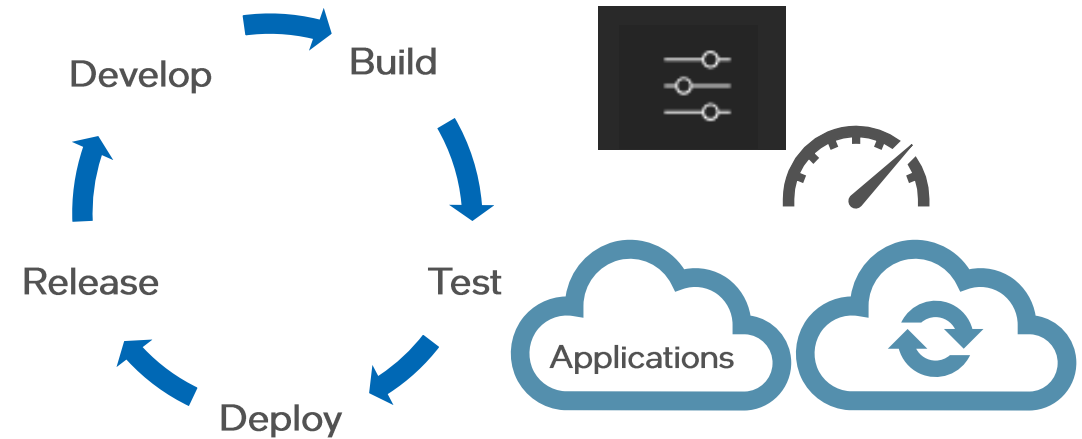
Application Decomposition

Containerized	Microservices

Microservice B

Microservice A   Microservice C

Life Cycle Automation

CI/CD/CD	Cloud Managed

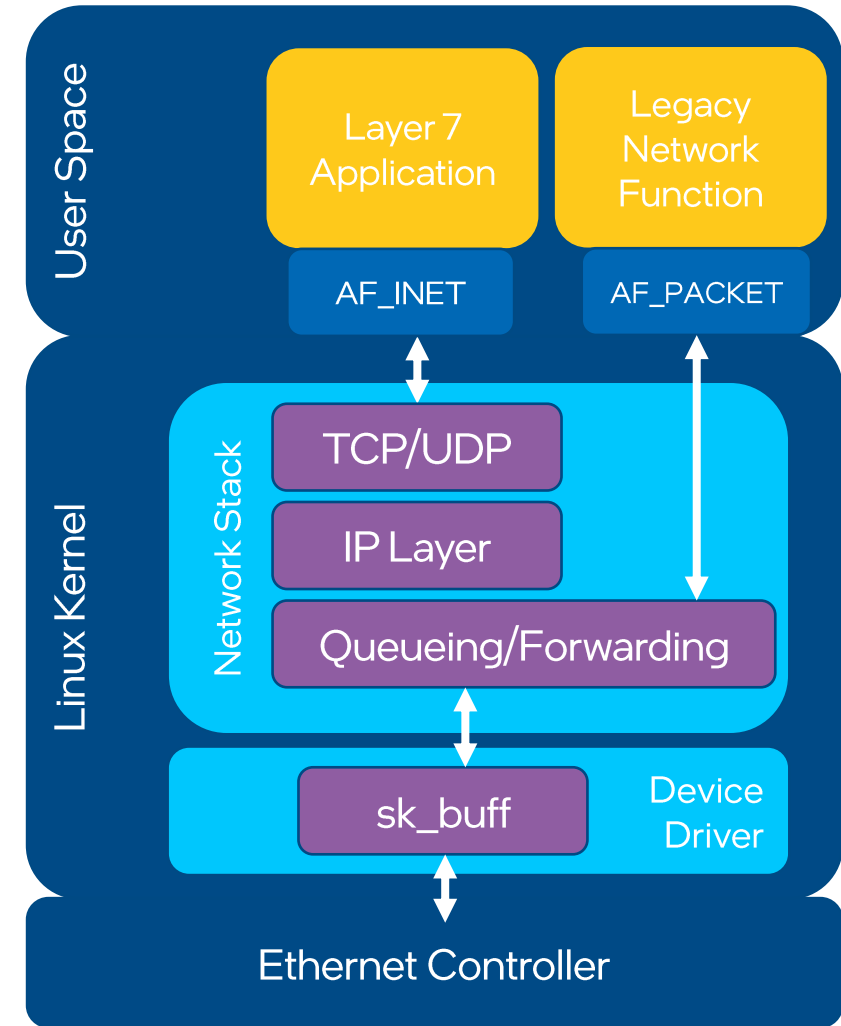Develop → Build → Test → Deploy → Release → Develop

Applications

Cloud Native Agile Approach Enables Innovation And Services Velocity

The Linux Networking Stack

- Sockets remain the main interface to all applications

- Linux Network Stack has evolved over many years

- Designed for features, compatibility, stability; not performance or custom use cases

Data Plane Development Kit

DPDK Scales on Intel® processor family from Atom to Xeon

- Accesses all devices directly from Linux User Space
- Software Optimization Techniques
- Framework abstracting application from platform
- Sample application code to showcase key processing capabilities
- Software Libraries to accelerate many packet processing needs

User Space

DPDK Application

NIC PMD

Layer 7 Application

AF_INET

Legacy Network Function

AF_PACKET

Linux Kernel

Network Stack

TCP/UDP

IP Layer

Queueing/Forwarding

sk_buff

Device Driver

Ethernet Controller
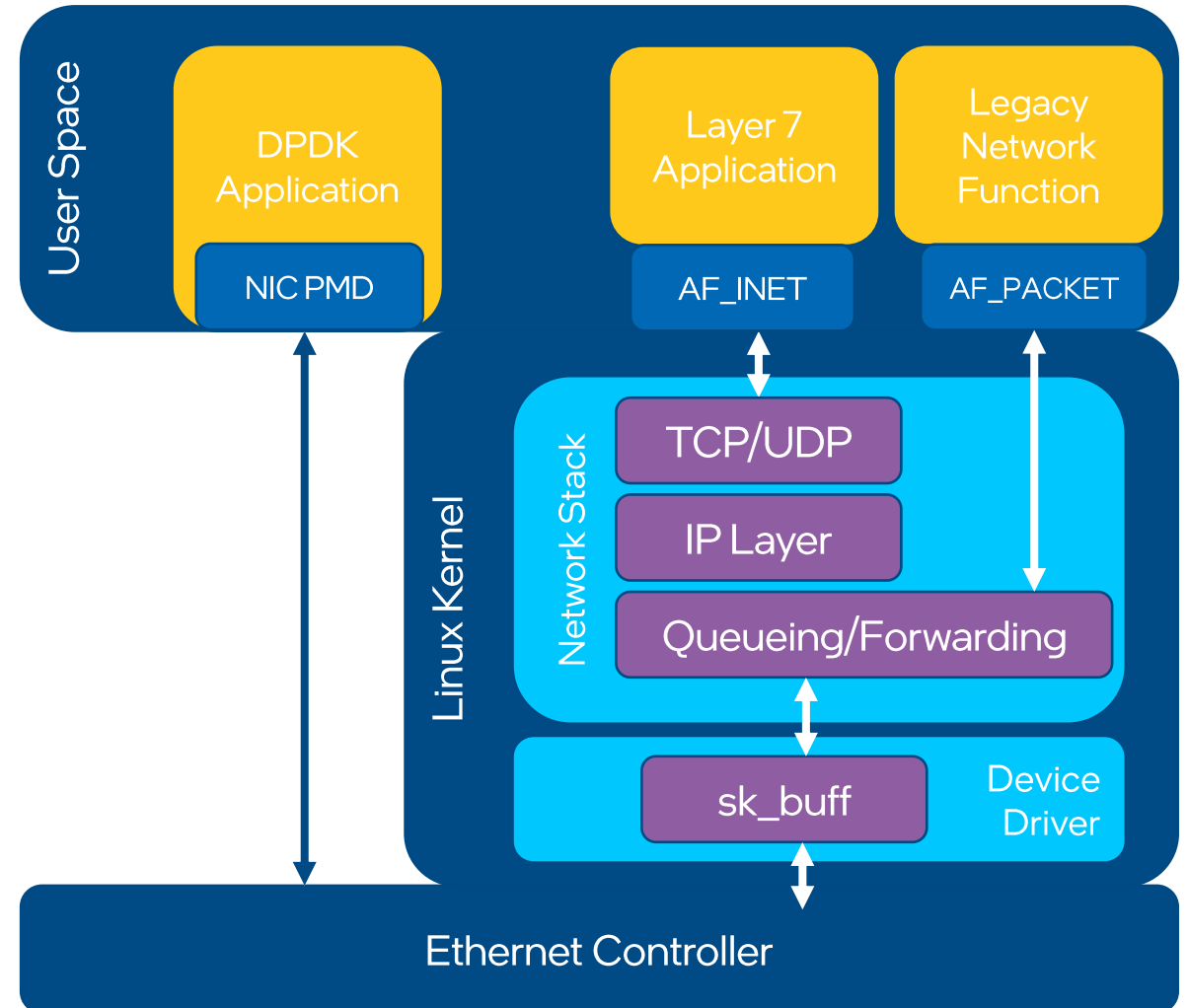
eBPF: Extended Berkeley Packet Filter

- Validated small program

- Executed for each packet

- Can be executed in many places

- Different actions and function in different places

- What eBPF can do depends on place it is executed

- Multiple eBPF programs can touch one packet
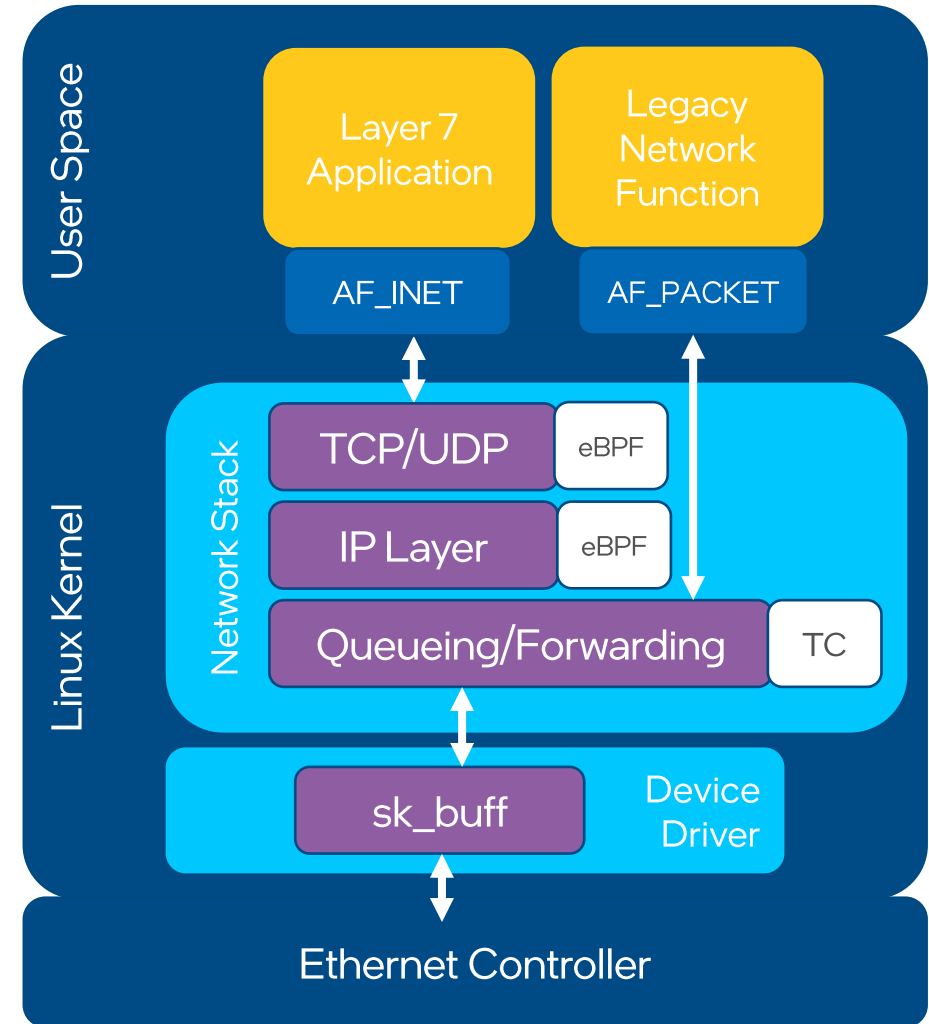
XDP: eXpress Data Path

- XDP = eBPF executed in the NIC driver

- Needs driver support; multiple vendors now supporting

- Actions: drop, pass, Tx or redirect

- Not a lot of information: L2

- Not executed on Tx from stack (future support)
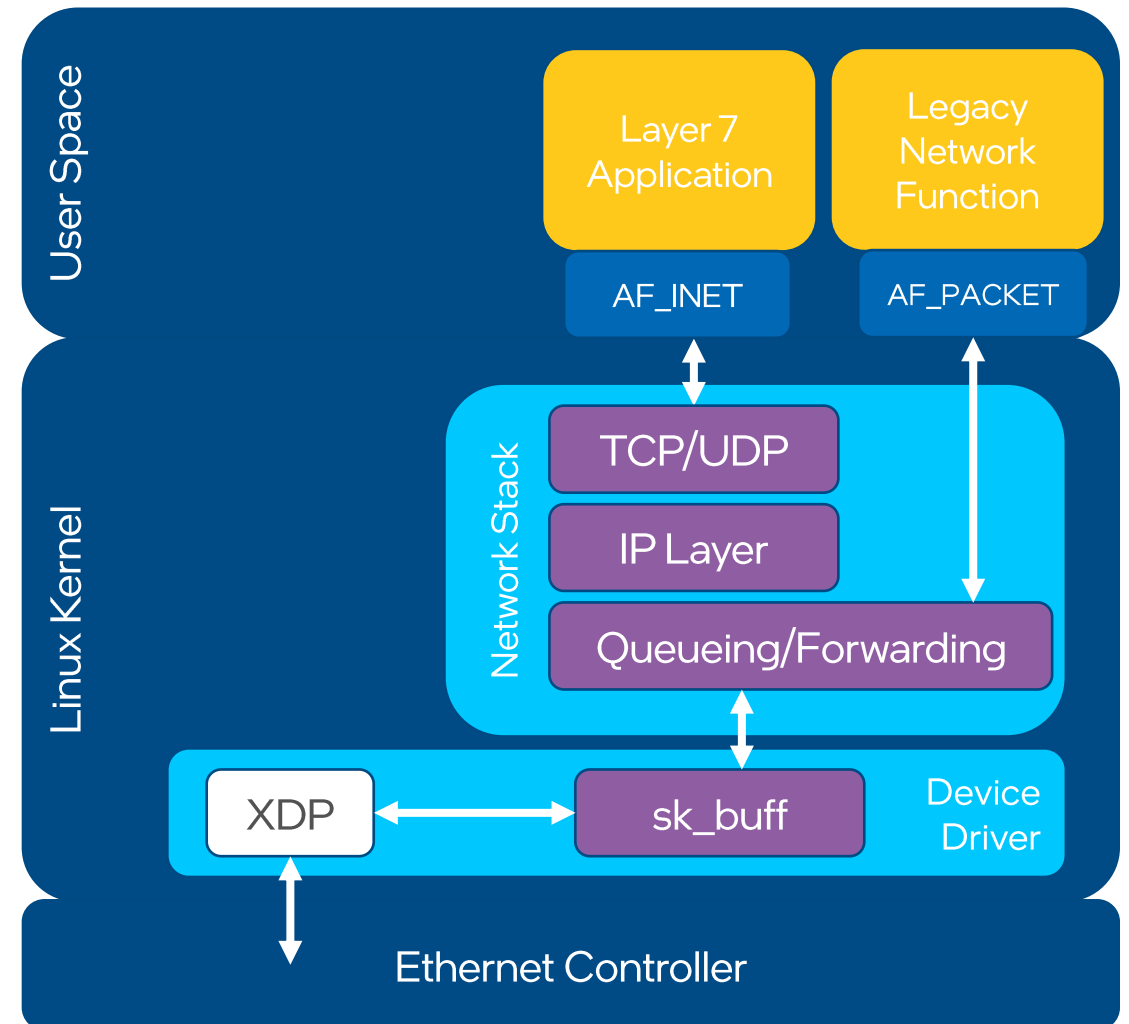
AF_XDP

- Fast packet delivery to user space

- Just a socket

- Can use poll() / select() as usual
  - Kernel and app driven from a single core
 - No infrastructure cores needed

- Libbpf for user space helpers
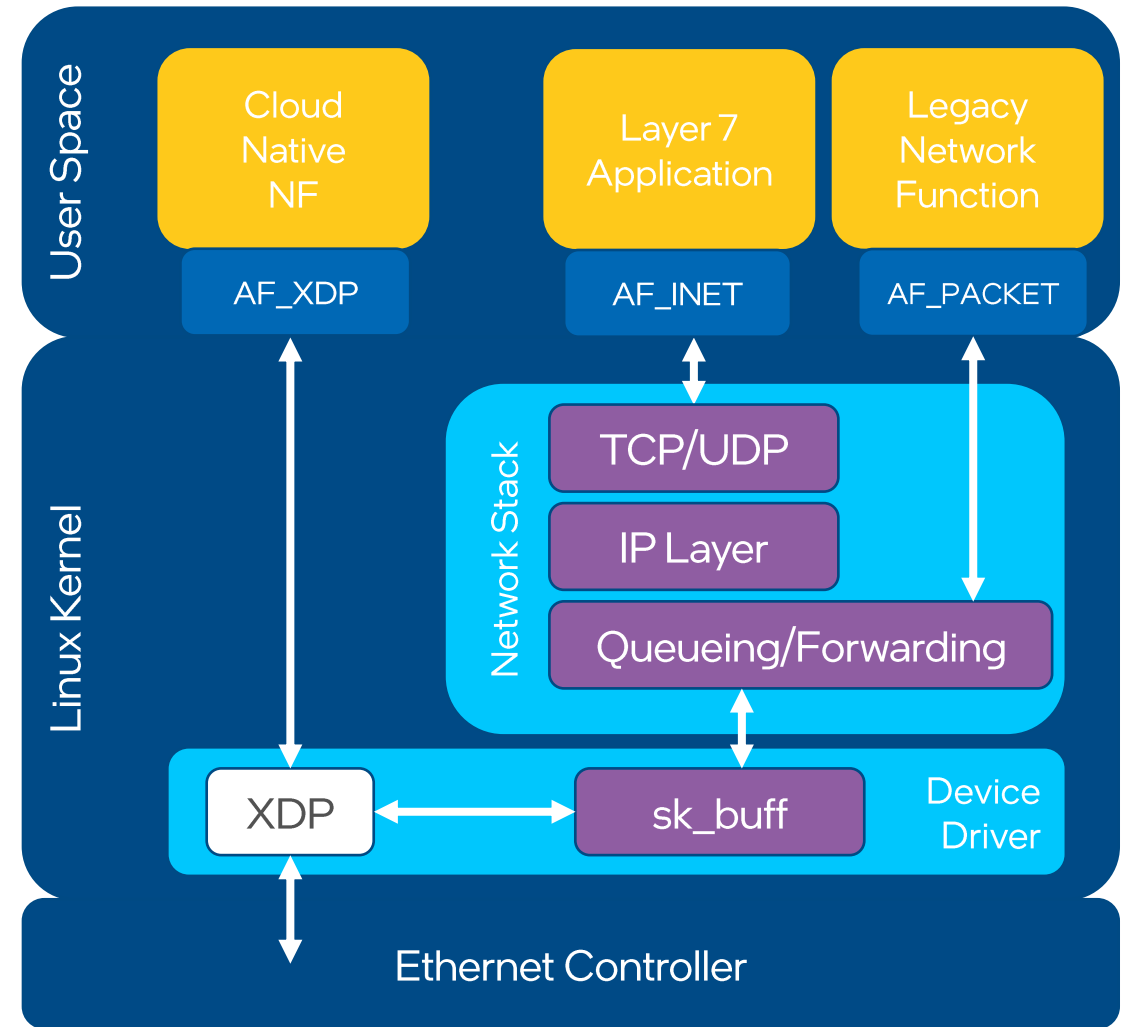
AF_XDP Zero Copy

- Very fast: zero-copy, batching, lockless

- XSKMAP for packet steering

- Since 4.18, August 2018. Supported on Intel® Ethernet 700/800 Series
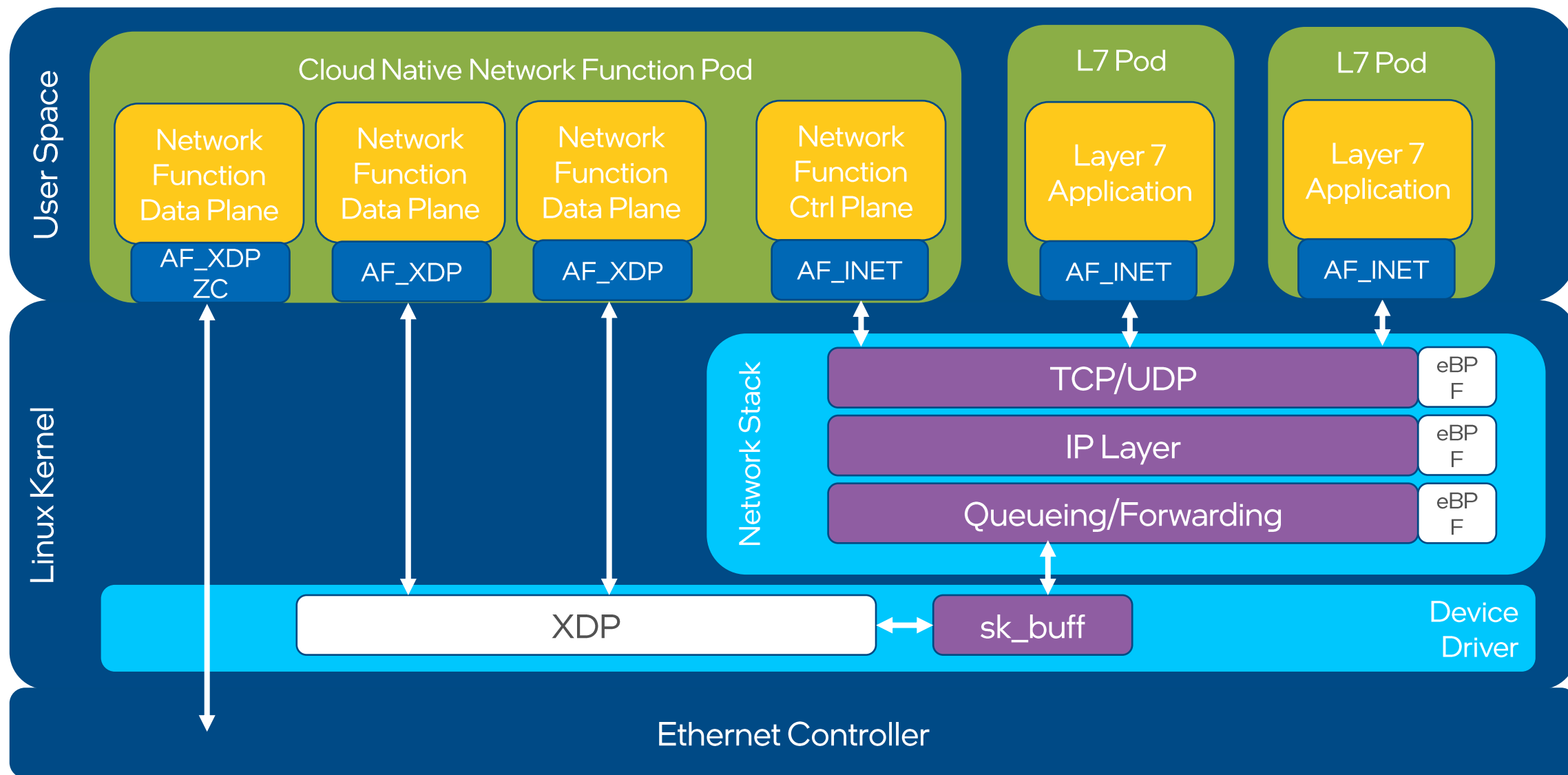
- DPDK PMD in 19.05 release
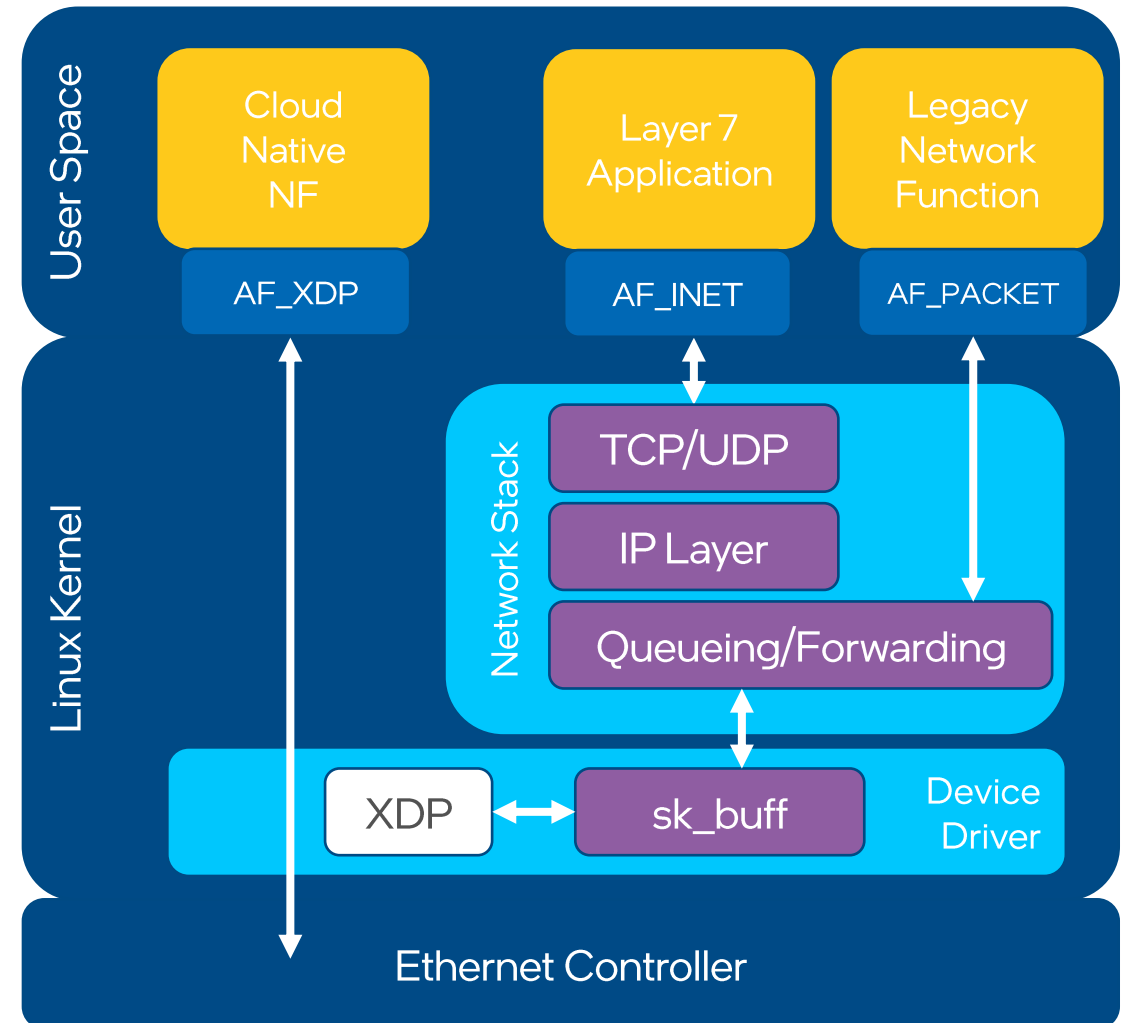
Use Case: Bifurcation

- ARP/ICMP/PTP etc usually handled by kernel

- Bad performance to go to user-space first then back to kernel

- Use XDP to bifurcate; works with or without hardware support
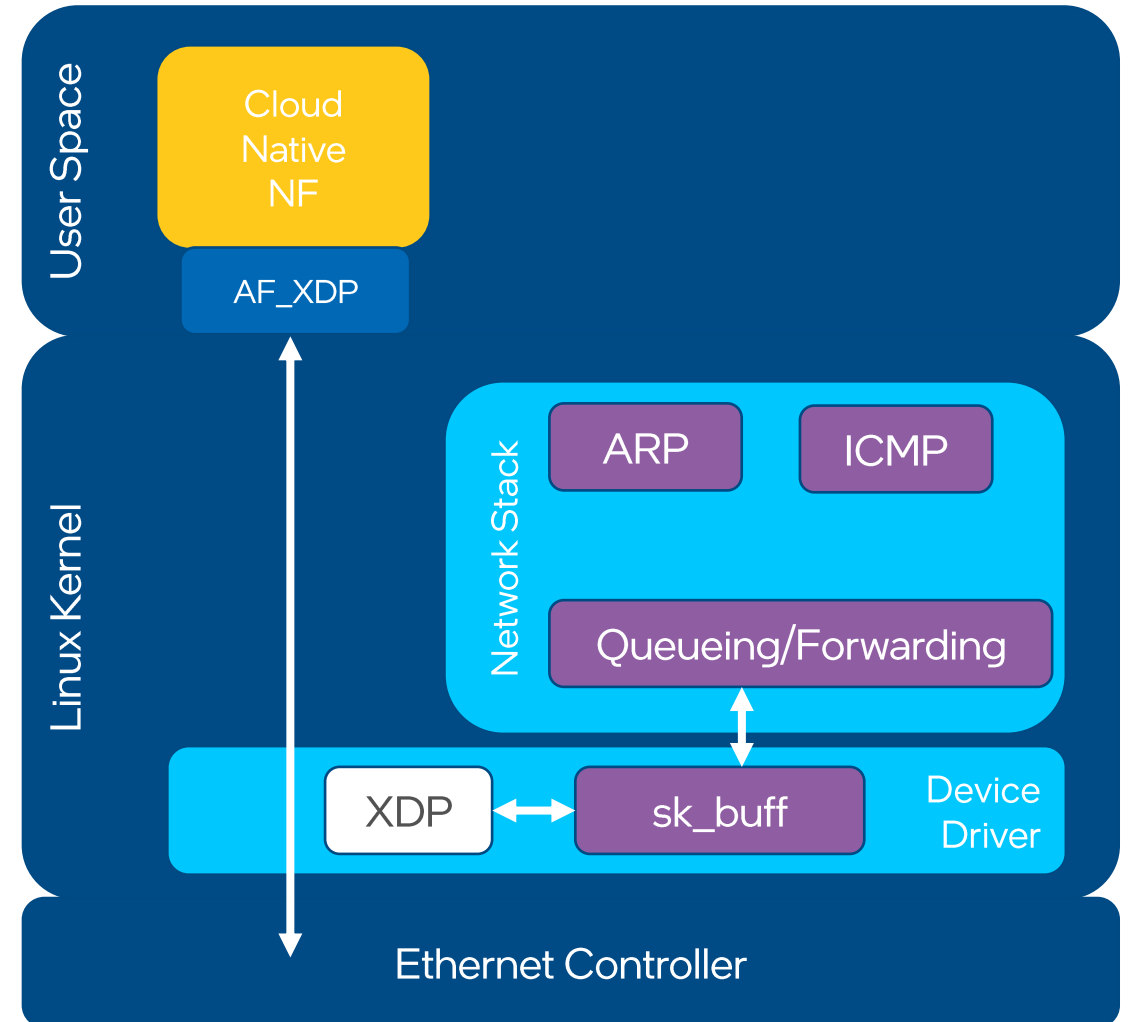
Use Case: Cloud Native Network Functions

Kubernetes and Linux Control

- For broadest use of Cloud Native Network Functions we need to work within existing security paradigms, do not want to mandate use of privileged pods

- Need a CNI plugin to provision the netdev when the pod is created; these are stateless, hence need a device plugin as well

- Currently an AF_XDP socket is tied to one netdev, which cannot currently be subdivided (containers cannot share a netdev)
  - Need a way of subdividing the netdev to allow individual queue assignment; currently being worked within the Linux kernel community

- Multus can allow multiple CNIs to be active

Kubernetes Support for Cloud Native Network Functions
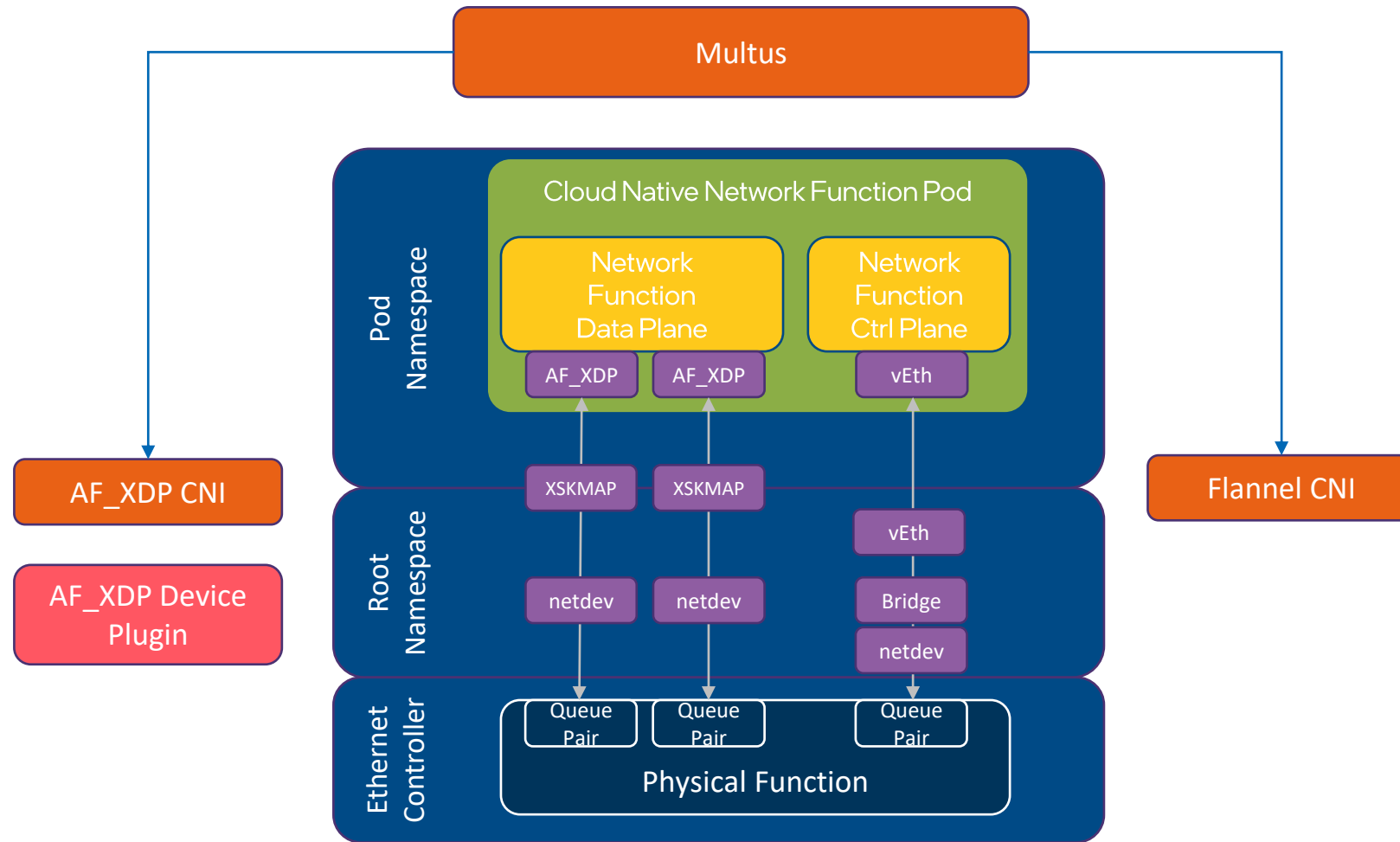
Addressing Gaps in Cloud Native Orchestration

Technology Gaps	Addressed By	
K8s Networking	Multiple Network for CNF	MULTUS
Packet Processing	High Performance E-W	CNI USERSPACE CNI DPDK
	High Performance N-S	CNI SR-IOV DPDK
	Dynamic Admission Controller	CNI Network Resource Injectors
	OVS-DPDK SDN Control	CNI KUBE-OVN
	Portable Data Plane N-S	CNI AF_XDP CNI
	HA Networking	BOND-CNI
	SR-IOV Metrics	SR-IOV Metrics Node Exporter
Telemetry	Scheduling per Telemetry	Telemetry Aware Scheduler
	Platform Telemetry	Collectd/Telegraf
Deployment	Deployment Playbook	

Addressing Gaps in Cloud Native Orchestration cont'd

Technology Gaps	Addressed By	
Resource Management	Platform discovery	Node Feature Discovery (Intel® AVX; SR-IOV; etc.)
	CPU pinning/isolation	CRI-RM
	Dynamic Huge Page	Huge page support for K8s
	Manage Devices	Device Plugins: (QAT, GPU, FPGA, VPU, SGX, DSA, SR-IOV)
	Set NUMA Alignment	Topology Manager (NUMA)
	Advanced NUMA Alignment	Devices, storage, memory, compute: CRI-RM
	SST & RDT (LLC/MBM)	Power operator with Native CPU Manager for K8s
	Next Gen holistic node resource management	CRI-RM

Summary

Cloud native is an application development approach that uses cloud computing delivery models

Working with existing and evolving Linux and Kubernetes mechanisms, we can achieve a balance of performance and abstraction for Cloud Native Network Functions

Intel supports closing the gaps in Cloud Native for service providers and delivers platforms that enable flexibility, agility and performance optimization that are foundational in modern network deployment

Find Out More

LEARN

About Kubernetes Networking Technologies
https://networkbuilders.intel.com/network-technologies/container-experience-kits

About Network Transformation Solutions, please click on:
https://networkbuilders.intel.com/network-technologies/network-transformation-exp-kits

About the Containerized 5G Core, please click on:
https://www.intel.com/content/www/us/en/communications/why-containers-and-cloud-native-functions-paper.html

EXPLORE

Free foundational 5G training from Intel® Network Academy*
https://www.coursera.org/learn/network-transformation-101

Container Bare Metal Kubernetes Reference Architecture
Container Bare Metal for 2nd Generation Intel® Xeon® Scalable Processor Reference Architecture

ENGAGE

Please contact your Intel representative

Links are provided in the attachments tab below your webinar screen

* Registration required.

Questions?

Xiaojun (Shawn) Li, Sales Director, Next Wave OEM & eODM

Xiaojun.Li@intel.com

Brian Skerry, Principal Engineer

Brian.J.Skerry@intel.com