Intel® Network Builders Insights Series

High Speed Packet Processing with the Data Plane Development Kit (DPDK)

- Tim O'Driscoll, Software Product Manager

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.

Data Plane Development Kit (DPDK)

DPDK
DATA PLANE DEVELOPMENT KIT

Intel Optimized

intel. XEON
intel. ATOM
intel. AGILEX
intel. eASIC
intel. OPTANE MEMORY

Intel® Ethernet Controller

Intel® QuickAssist Technology

Open Source, Multi-Vendor [1]

arm
Chelsio Communications Accelerate
BROADCOM
NXP
MARVELL
AMD
HISILICON
SOLARFLARE®
NVIDIA

DPDK
DATA PLANE DEVELOPMENT KIT

The Data Plane Development Kit (DPDK) is a set of software libraries and drivers for accelerating packet processing workloads on COTS hardware platforms.

Commercially Proven [2]

ZTE
NTT
AT&T
HUAWEI
Alibaba Group
cisco
Baidu 百度
Microsoft
f5
nuagenetworks From Nokia
NEC
COMCAST
vmware
ERICSSON

Open Source Ecosystem [3]

OvS Open vSwitch
.io
F-Stack
redhat.
mTCP
fedora
SPDK
CentOS®
ubuntu
tungstenfabric
vmware ESXi

1. Source: DPDK contributors
2. Source: DPDK project members and DPDK Summit presenters
3. Source: Open source projects using DPDK

The Network Transformation Journey

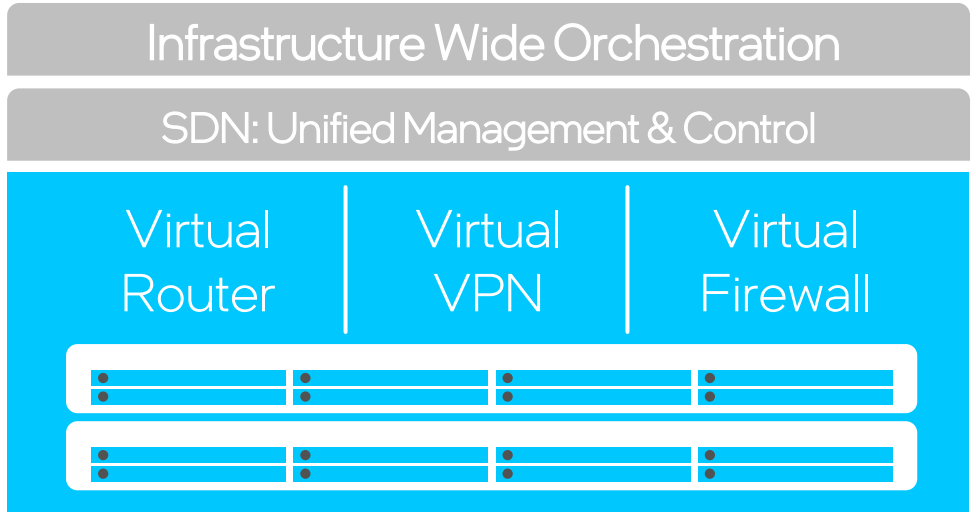
Driving Network Scale & Agility

Yesterday: Purpose-Built

Siloed Management & Control

Transforming:
Architecture
Value Chain
Business Process

Modern: Built for Multi-Purpose

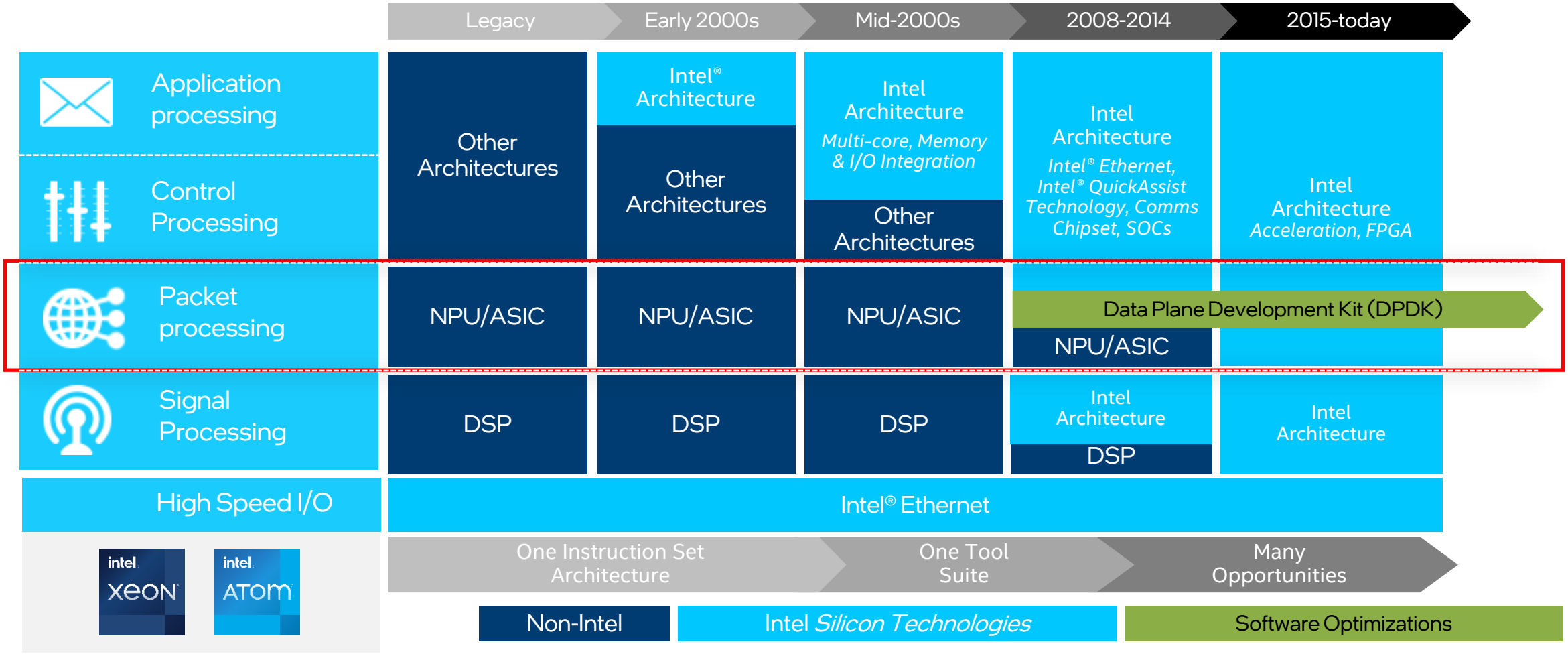


Proprietary hardware and software with an optimized data plane

Open, standard platform serving diverse workloads

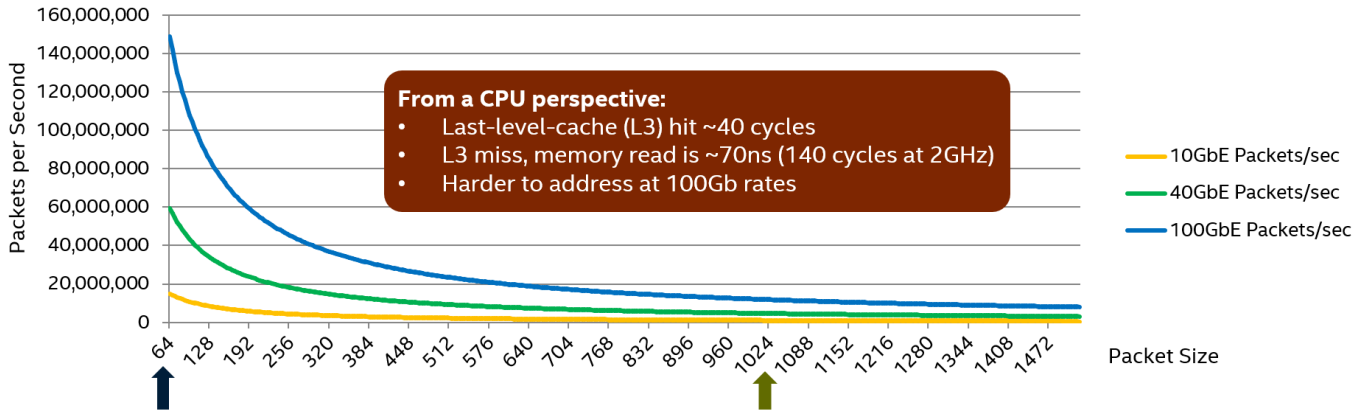
Over a Decade of Investment in Workload Convergence

Hardware & Software Optimizations Allow the Migration of all Network Services to IA Platforms



High Performance Challenges

Cycles count



Packet Size	64 Bytes
40G packets/second	59.5 million each way
Packet arrival interval	16.8 ns
2 GHz clock cycles/packet	33 cycles

Typical Network Infrastructure Packet Size

Packet Size	1024 Bytes
40G packets/second	4.8 million each way
Packet arrival interval	208.8 ns
2 GHz clock cycles/packet	417 cycles

Typical Server Packet Size

The system can't keep up with the interrupts for packet Rx:

- Switch from an interrupt-driven to a polled-mode driver.

Linux scheduler high overhead for task switches:

- Bind a single software thread to a logical core.

Memory and PCIe access is slow compared to CPU operations:

- Process a bunch of packets during each software iteration and amortize the access cost over multiple packets.

Data doesn't seem to be near the CPU when it needs to be:

- For memory access, use HW or SW controlled prefetching. For PCIe access, use Data Direct IO to write data directly into cache.

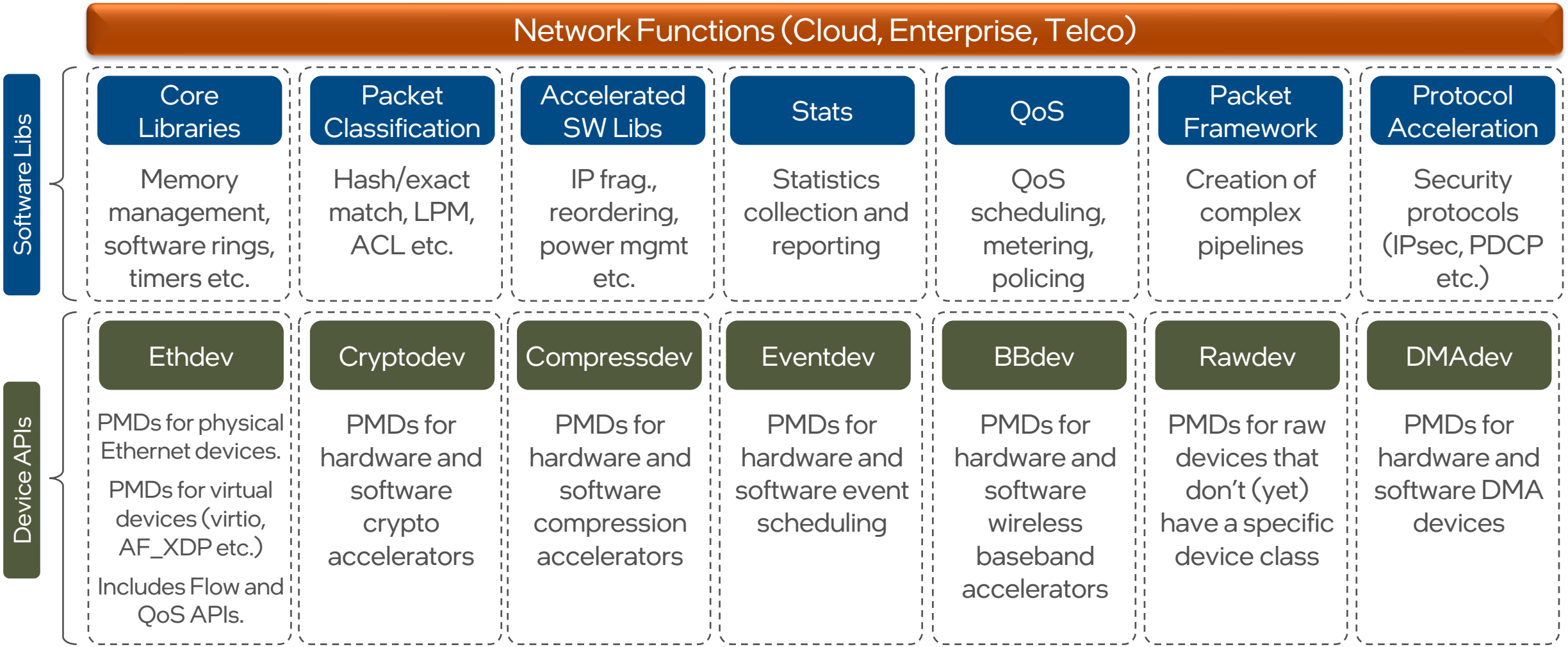
Access to shared data structures is a bottleneck:

- Use access schemes that reduce the amount of sharing (e.g. lockless queues for message passing).

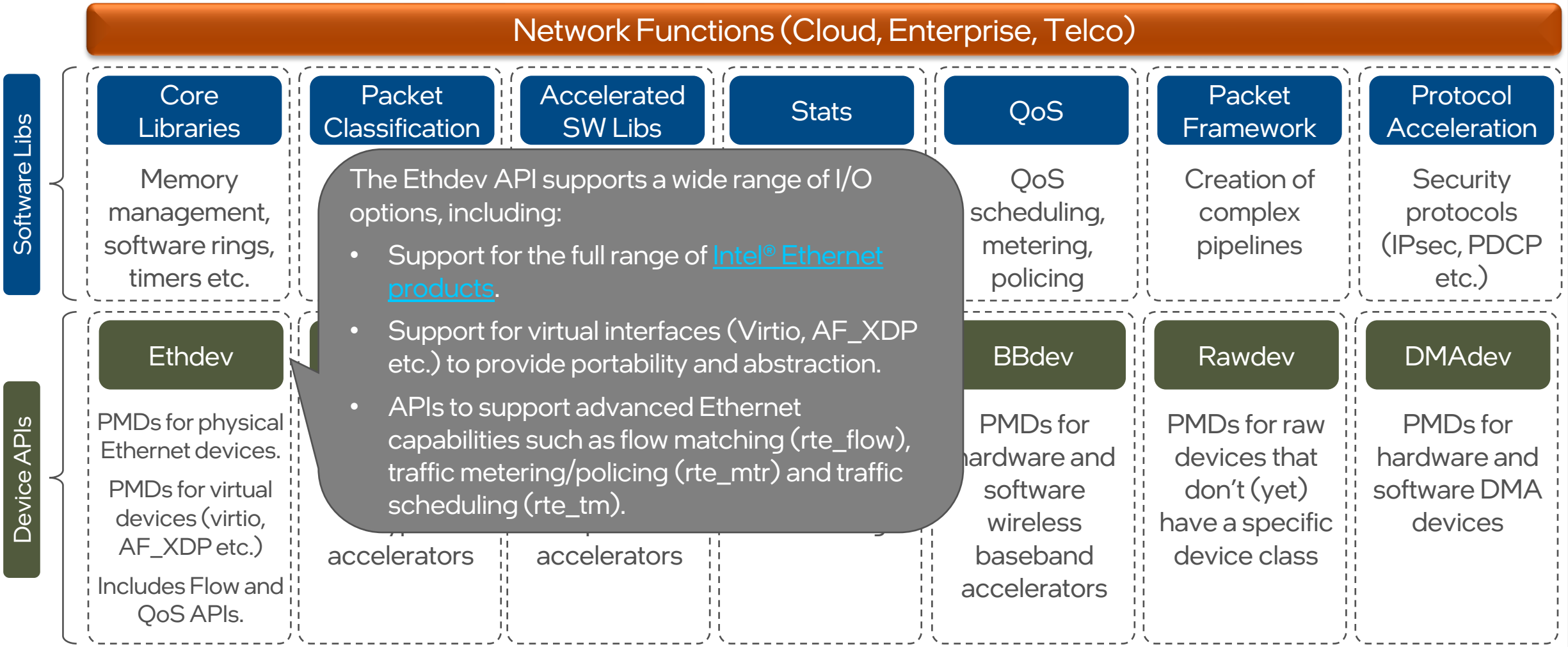
Page tables are constantly evicted (DTLB Thrashing):

- Allow Linux to use Huge Pages (2MB, 1GB)

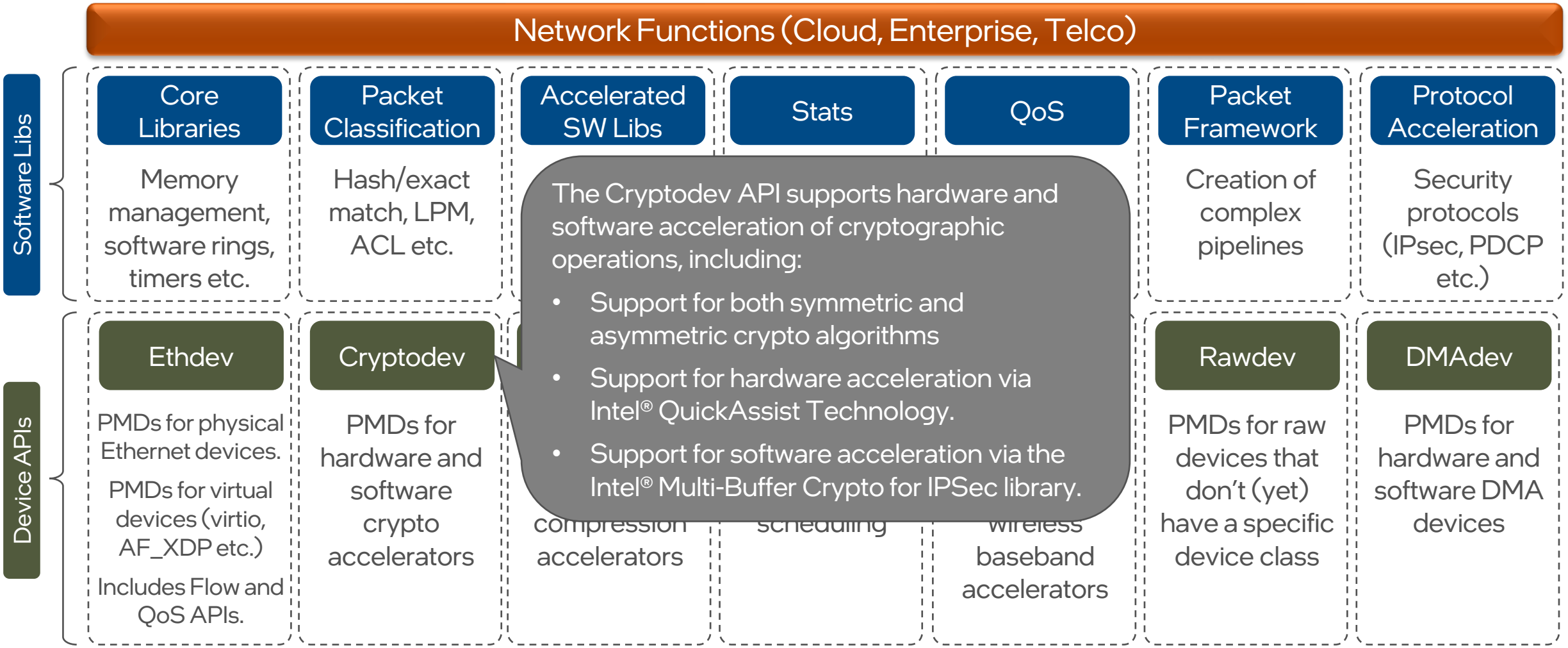
DPDK Scope


Network Functions (Cloud, Enterprise, Telco)
Software Libs
Core Libraries
Memory management, software rings, timers etc.
Packet Classification
Hash/exact match, LPM, ACL etc.
Accelerated SW Libs
IP frag., reordering, power mgmt etc.
Stats
Statistics collection and reporting
QoS
QoS scheduling, metering, policing
Packet Framework
Creation of complex pipelines
Protocol Acceleration
Security protocols (IPsec, PDCP etc.)
Device APIs
Ethdev
PMDs for physical Ethernet devices.
PMDs for virtual devices (virtio, AF_XDP etc.)
Includes Flow and QoS APIs.
Cryptodev
PMDs for hardware and software crypto accelerators
Compressdev
PMDs for hardware and software compression accelerators
Eventdev
PMDs for hardware and software event scheduling
BBdev
PMDs for hardware and software wireless baseband accelerators
Rawdev
PMDs for raw devices that don't (yet) have a specific device class
DMAdev
PMDs for hardware and software DMA devices

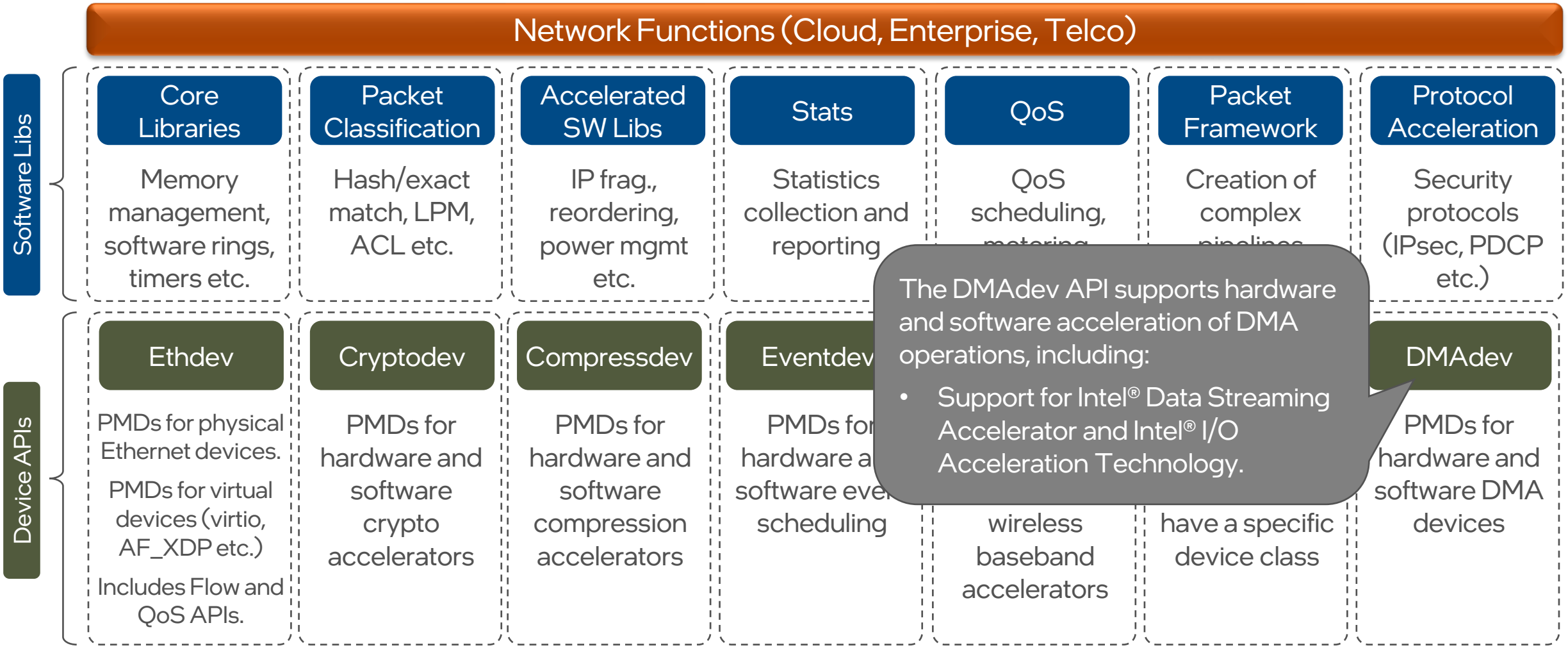
DPDK Scope


Network Functions (Cloud, Enterprise, Telco)
Software Libs
Core Libraries
Memory management, software rings, timers etc.
Packet Classification
Accelerated SW Libs
Stats
QoS
QoS scheduling, metering, policing
Packet Framework
Creation of complex pipelines
Protocol Acceleration
Security protocols (IPsec, PDCP etc.)
The Ethdev API supports a wide range of I/O options, including:
• Support for the full range of Intel® Ethernet products.
• Support for virtual interfaces (Virtio, AF_XDP etc.) to provide portability and abstraction.
• APIs to support advanced Ethernet capabilities such as flow matching (rte_flow), traffic metering/policing (rte_mtr) and traffic scheduling (rte_tm).
Device APIs
Ethdev
PMDs for physical Ethernet devices.
PMDs for virtual devices (virtio, AF_XDP etc.)
Includes Flow and QoS APIs.
accelerators
accelerators
BBdev
PMDs for hardware and software wireless baseband accelerators
Rawdev
PMDs for raw devices that don't (yet) have a specific device class
DMAdev
PMDs for hardware and software DMA devices

DPDK Scope

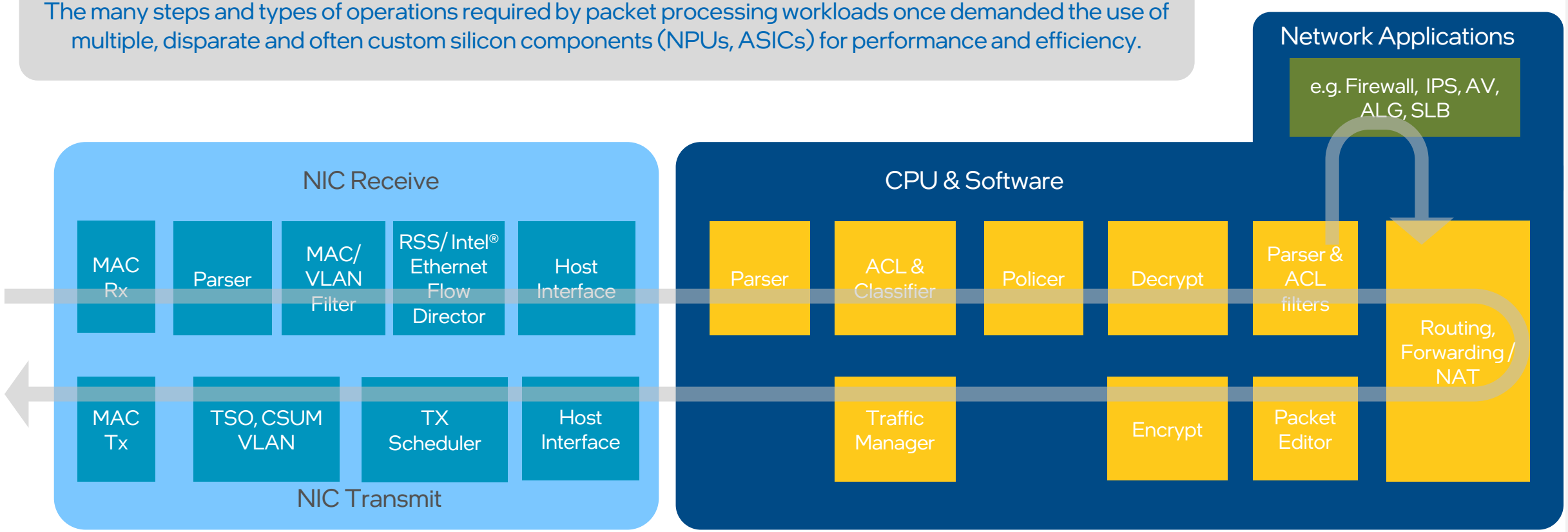

Network Functions (Cloud, Enterprise, Telco)
Software Libs
Core Libraries
Memory management, software rings, timers etc.
Packet Classification
Hash/exact match, LPM, ACL etc.
Accelerated SW Libs
Stats
QoS
Packet Framework
Creation of complex pipelines
Protocol Acceleration
Security protocols (IPsec, PDCP etc.)
The Cryptodev API supports hardware and software acceleration of cryptographic operations, including:
• Support for both symmetric and asymmetric crypto algorithms
• Support for hardware acceleration via Intel® QuickAssist Technology.
• Support for software acceleration via the Intel® Multi-Buffer Crypto for IPSec library.
Device APIs
Ethdev
PMDs for physical Ethernet devices.
PMDs for virtual devices (virtio, AF_XDP etc.)
Includes Flow and QoS APIs.
Cryptodev
PMDs for hardware and software crypto accelerators
compression accelerators
scheduling
wireless baseband accelerators
Rawdev
PMDs for raw devices that don't (yet) have a specific device class
DMAdev
PMDs for hardware and software DMA devices

DPDK Scope


Network Functions (Cloud, Enterprise, Telco)
Software Libs
Core Libraries
Memory management, software rings, timers etc.
Packet Classification
Hash/exact match, LPM, ACL etc.
Accelerated SW Libs
IP frag., reordering, power mgmt etc.
Stats
Statistics collection and reporting
QoS
QoS scheduling,
Packet Framework
Creation of complex
Protocol Acceleration
Security protocols (IPsec, PDCP etc.)
Device APIs
Ethdev
PMDs for physical Ethernet devices.
PMDs for virtual devices (virtio, AF_XDP etc.)
Includes Flow and QoS APIs.
Cryptodev
PMDs for hardware and software crypto accelerators
Compressdev
PMDs for hardware and software compression accelerators
Eventdev
PMDs for hardware a
software eve
scheduling
wireless baseband accelerators
have a specific device class
DMAdev
PMDs for hardware and software DMA devices
The DMAdev API supports hardware and software acceleration of DMA operations, including:
• Support for Intel® Data Streaming Accelerator and Intel® I/O Acceleration Technology.

Example DPDK Application

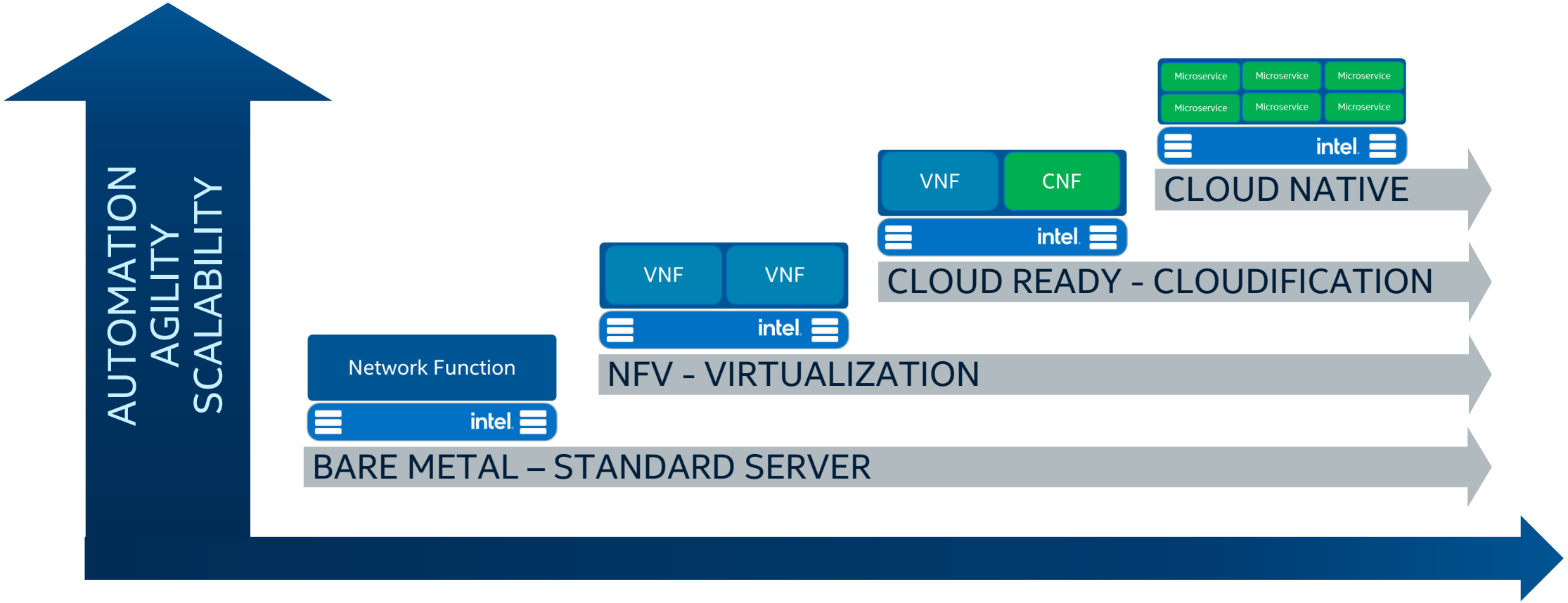
The many steps and types of operations required by packet processing workloads once demanded the use of multiple, disparate and often custom silicon components (NPUs, ASICs) for performance and efficiency.



Today, the right combination of platform, ingredients, software optimization and ecosystem support can enhance even these demanding packet workloads.

NIC | DPDK | Application

Networking Evolution



DPDK continues to evolve to meet the needs of the networking industry.

Networking Evolution - DPDK

Key DPDK Enhancements

Virtual/Cloud Interfaces
- Virtio (inc. virtio 1.1, virtio-user, vDPA)
- AF_XDP
- VMware VMXNET3
- Amazon ENA
- Microsoft NETVSC
- Google VNIC (in progress)

Transparent Acceleration
- Acceleration APIs support both hardware and software acceleration: cryptodev, compressdev, eventdev, BBdev etc.

Flexibility/Ease of Use
- Dynamic memory management
- Live migration
- K8S integration

And many more …

References

High Performance Cloud-Ready Virtual Network Functions with DPDK Device Abstraction

Validating DPDK Application Portability in Multi-cloud/Hybrid-cloud Environments

Virtio-networking series

Making high performance networking applications work on hybrid clouds

Breaking Cloud Native Network Performance Barriers

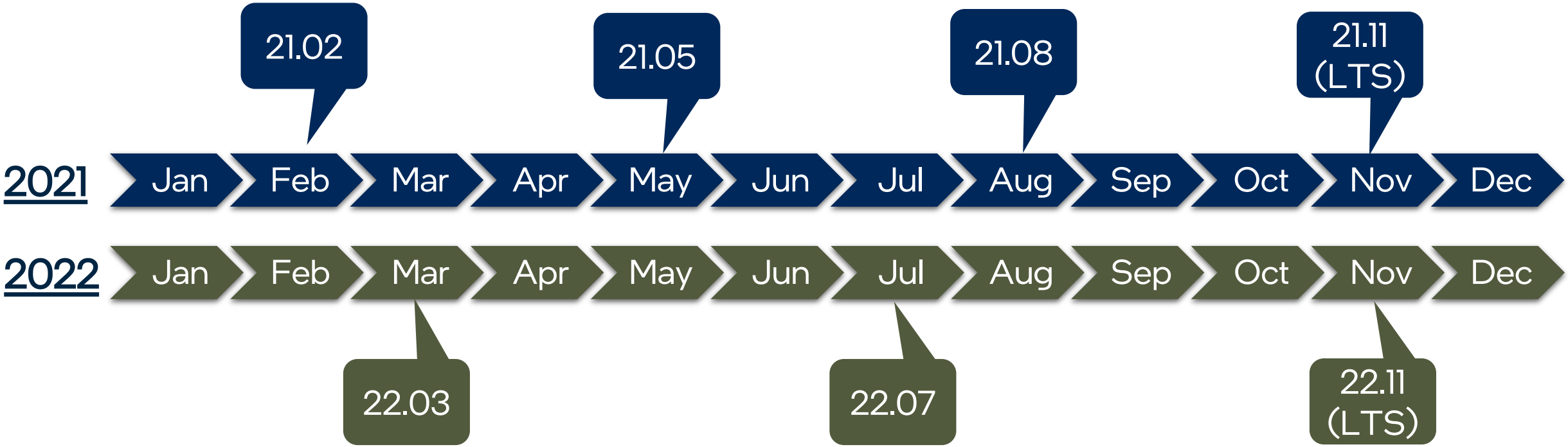
Getting Started with DPDK on AWS EC2

DPDK on Azure

DPDK in an Azure Linux VM

And many more …

DPDK Release Cadence



- Change from 4 to 3 releases per year in 2022.
- Long-Term Support (LTS) provides a stable release maintained for 2 years:
  - Provides a stable target on which to base applications or packages.
  - Bug fixes that do not change the ABI will be back-ported.
  - In general, new features will not be back-ported.

Case Studies

DPDK
DATA PLANE DEVELOPMENT KIT

vRAN: The Next Step in Network Transformation

ALTIOSTAR
Leading Network Transformation

WIND
AN INTEL COMPANY

DELL

amdocs

*"Based on the **Data Plane Development Kit (DPDK)**, AVS (Accelerated Virtual Switch) achieves line rate virtual switching performance using fewer CPU cores than any other virtual switch. The Altiostar vRAN solution takes full advantage of multi-socket and multi-core processors and **DPDK** support for maximum use of underlying hardware resources*

5G Core Network UPF

*"Together, Intel® processors, Intel® network adapters, **DPDK**, and other relevant technologies form a platform that reduces device complexity and improves network utilization to meet telecom carriers' requirements for the flexibility, scalability, reliability, and cost efficiency "*

ZTE

NEC's UPF Maximizes 5G Value

*"NEC uses Intel's technologies of **Data Plane Development Kit (DPDK)** and Dynamic Device Personalization (DDP)."*
"Leveraging the benefits of cloud native UPF allows flexible and rapid deployment scenarios based on traffic and network characteristics."

NEC

Metaswitch SBC 26% Improvement

*"Benchmark tests were configured using the **Data Plane Development Kit (DPDK)** to improve packet processing performance and optimize the system for maximum advanced media capacity."*

metaswitch

Clavister NetShield Firewall Scales Linearly

*"Clavister utilized the open source **Data Plane Development Kit (DPDK)**, a set of software libraries and drivers in NetShield Virtual's scalable data plane. DPDK facilitates high-performance data throughput in an Intel® architecture-based server"*

clavister

Further Information

Documentation

- All documentation
- Getting Started Guide for Linux
- Programmer's Guide
- How-To Guides
- Network Interface Controller Drivers
- Crypto Device Drivers
- Compression Device Drivers
- Baseband Device Drivers
- Event Device Drivers
- DMA Device Drivers
- Release Notes
- Sample Applications User Guides
- Testpmd Application User Guide
- And many more …

DPDK Summits

- Upcoming events
- Past events

Testing/CI

- DPDK Testing/CI overview
- DPDK Community Lab
- DPDK Test Suite (DTS)
- DPDK Performance Reports

Training

- DPDK Cookbook
- Intel® Network Builders University – DPDK

Misc.

- DPDK Whitepaper
- Intel® Network Builders - Network Transformation Experience Kits
- DPDK Roadmap

Questions?

Xiaojun (Shawn) Li, Sales Director, Next Wave OEM & eODM
Xiaojun.Li@intel.com

Tim O'Driscoll, Software Product Manager
Tim.odriscoll@intel.com

Join Us Next Time

July 13th @ 8am PDT

Intel® Network Builders Insights Series

Reference Implementations (RIs) - Free, Customer-Deployable Apps for Customer Use Cases

- Hassnaa Moustafa, Principal Engineer, Intel Corporation

intel