Network Platform on 3rd Gen Intel® Xeon® Scalable Platform

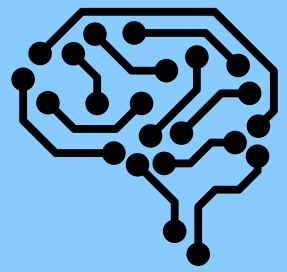
Tao Yang (杨涛)

Intel Corporation

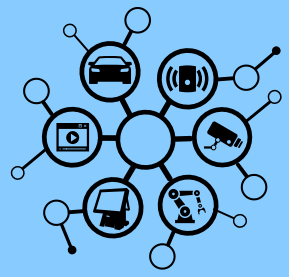
Key Industry Inflections

5G Network Transformation

Artificial Intelligence

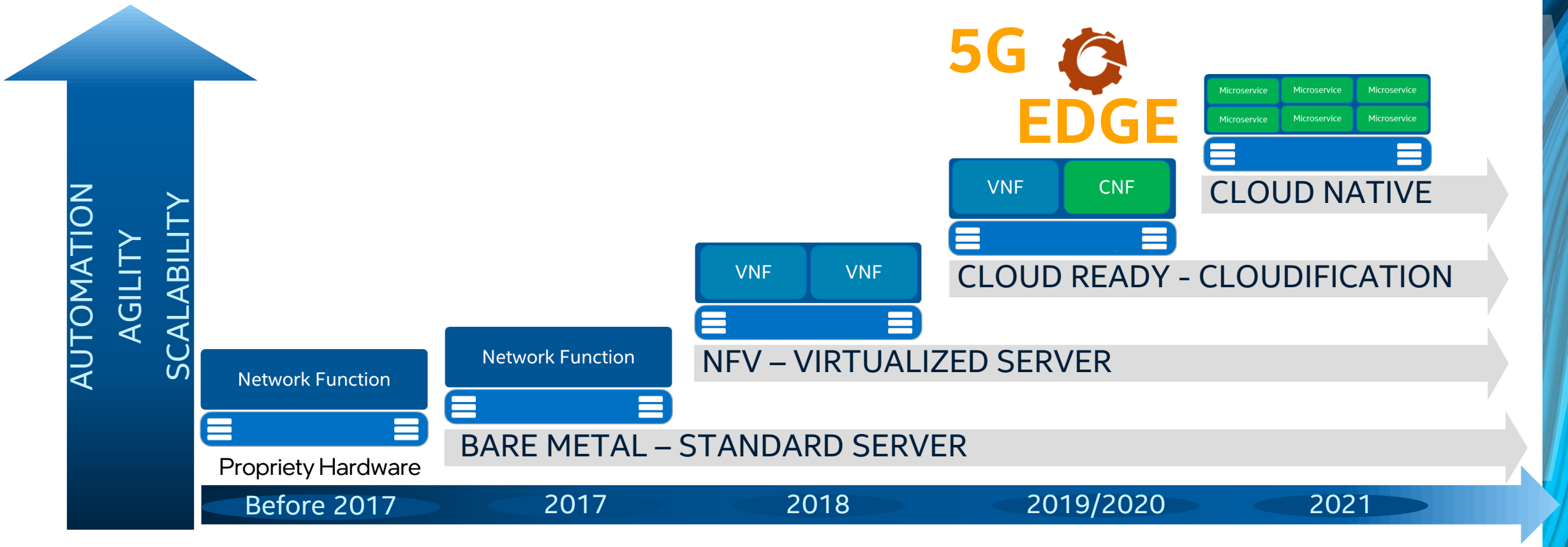


Intelligent Edge

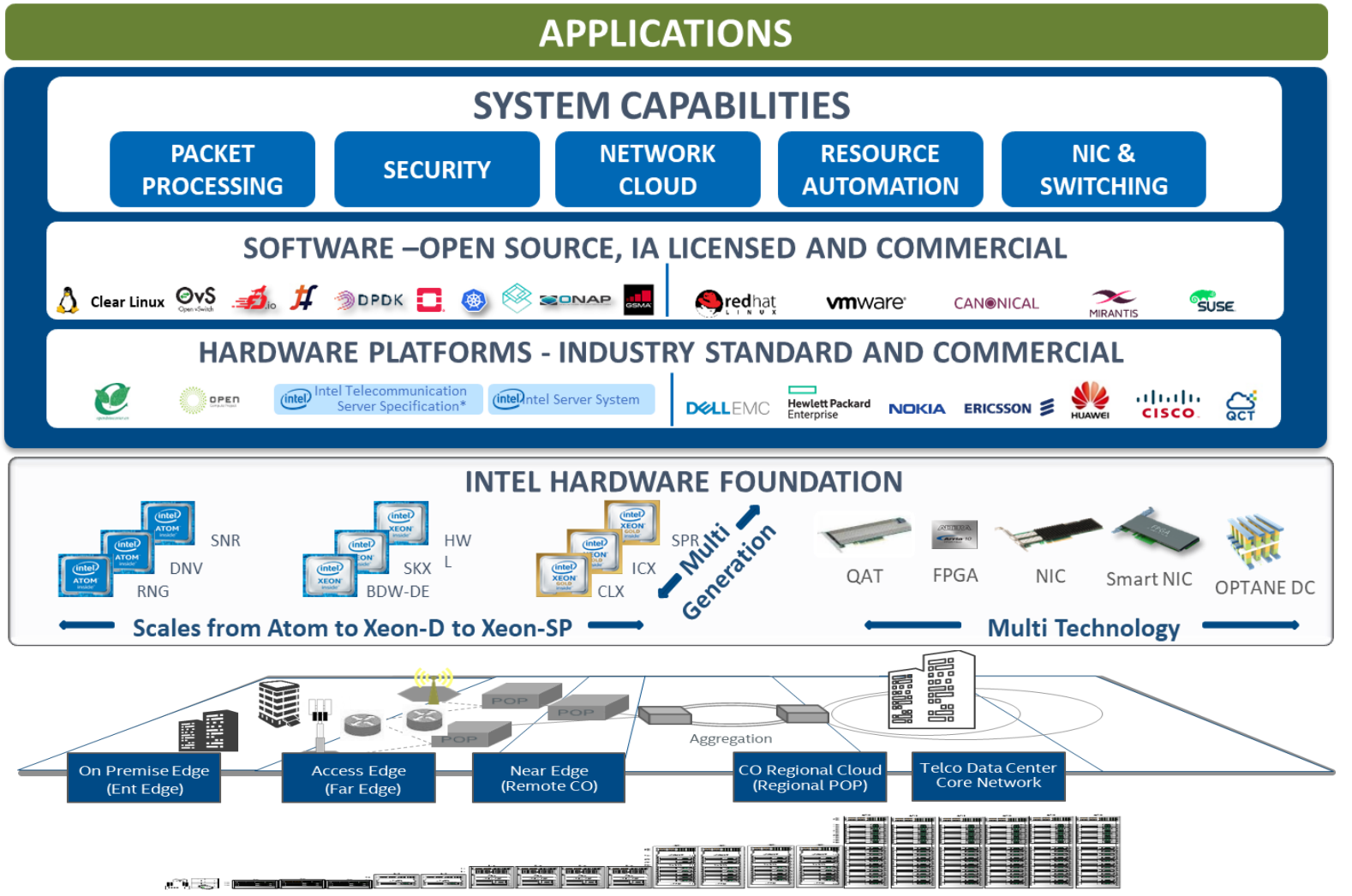


Cloudification of Everything

The Journey of Network Platform Transform



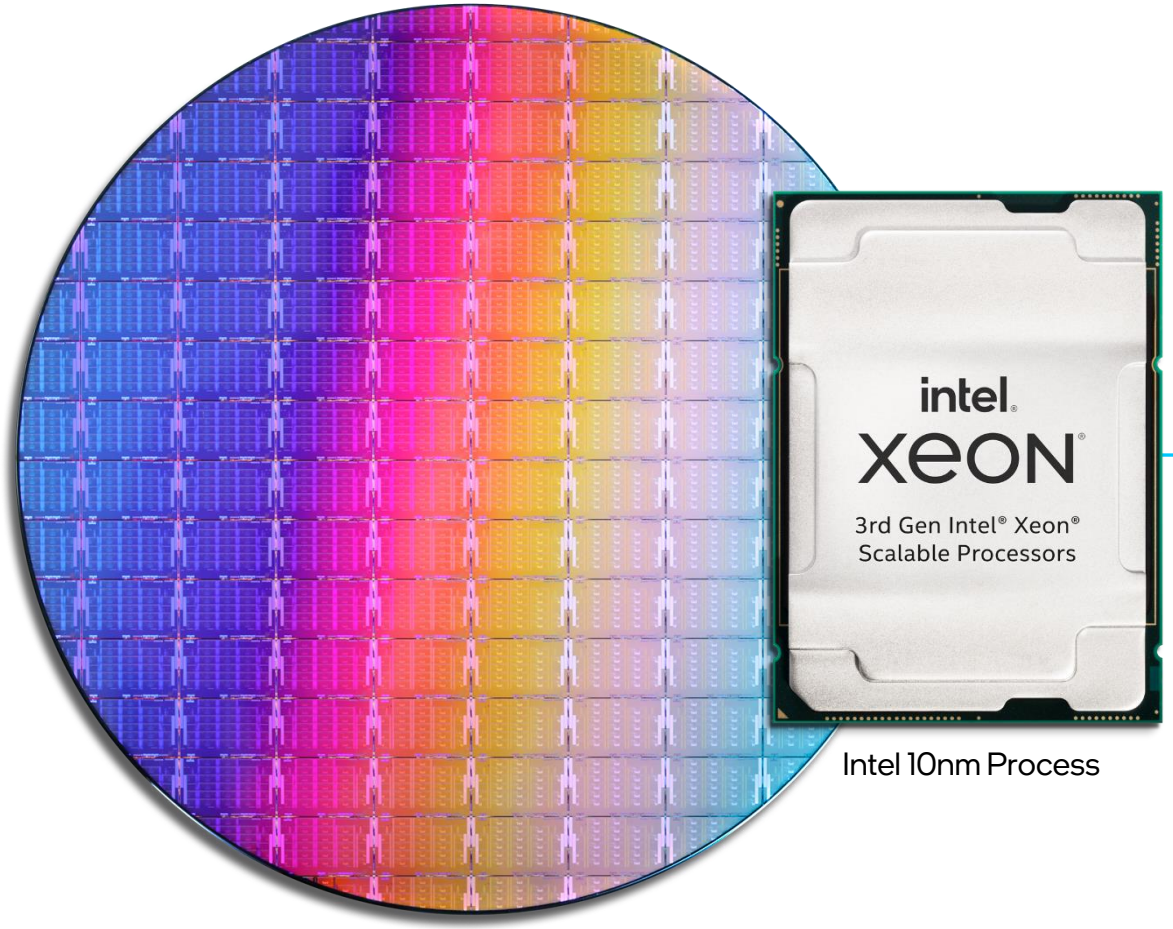
Network Platform on Intel Architecture



INTRODUCING

3rd Gen Intel® Xeon® Scalable processors

Performance made flexible



Intel 10nm Process

Up to 40 cores
per processor

20% IPC improvement
28 core, ISO Freq, ISO compiler

1.46x average performance increase
Geomean of Integer, Floating Point, Stream Triad, LINPACK
8380 vs. 8280

1.74x AI inference increase
8380 vs. 8280 BERT

2.65x average performance increase
vs. 5-year-old system
8380 vs. E5-2699v4

Performance varies by use, configuration and other factors. Configurations see appendix [1,3,5,55]

INTRODUCING

3rd Gen Intel® Xeon® Scalable processors

Performance made flexible

Only x86 data center processor with built-in AI & security solutions

Targeted for 1S-2S systems

Advanced security solutions

- Intel Software Guard Extensions
- Intel Crypto Acceleration
- Intel Total Memory Encryption
- Intel Platform Firmware Resilience

Scalable, flexible, customizable

- Intel Deep Learning Boost
- Intel Speed Select Technology
- Intel AVX-512
- oneAPI POWERED Optimized Software

Next-gen Xeon Scalable **Platform**

Up to 6TB	Up to 8CH	Up to 2.6X	Up to 64
System Memory Capacity (Per Socket) DRAM + PMem	DDR4-3200 2 DPC (Per Socket)	Memory Capacity Increase vs. 2nd Gen Xeon Scalable	Lanes PCI Express 4 (per Socket)

Breakthrough Data **Performance**

- Intel® Optane™ persistent memory 200 series
- Intel® Optane™ SSD P5800X series
- Intel® SSD D series

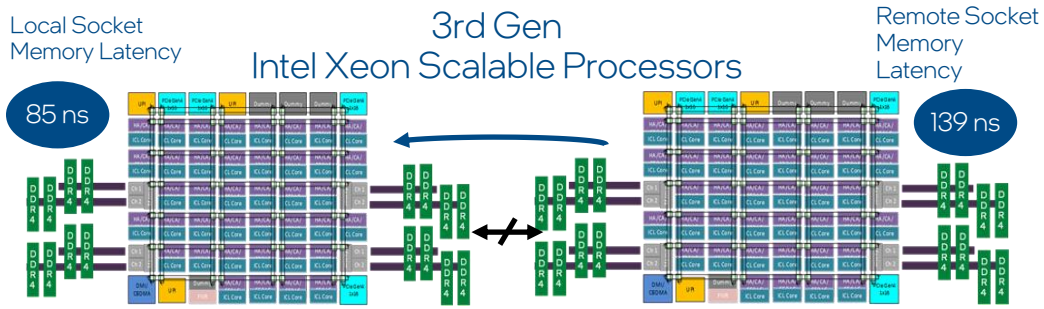
Faster, Flexible, Data **Scale**

- Intel® Ethernet 800 series network adapters
- Intel® Agilex™ FPGA solutions
- intel SELECT SOLUTIONS
- intel MARKET READY

Performance made flexible.

3rd Gen Intel® Xeon® Scalable Processors

Memory controller, memory latency and capacity



	Intel Xeon Platinum 8380 Processor (Ice Lake)	Intel Xeon Platinum 8280 Processor (Cascade Lake)
Memory Controller	On die – 8 ch	On Die – 6ch
Max DIMM Capability	2 DPC 3200/2933/2666 (SKU dependent)	1 DPC 2933/2 DPC 2666 (SKU dependent)
DRAM read latency local socket, ns	85	81
DRAM read latency (remote socket), ns	139	138
Max Memory Capacity per Socket	6TB (DDR+Pmem). 4TB (DDR)	4.5 TB (DDR+Pmem). 3TB (DDR)

3rd Gen Xeon Scalable provides lower latencies to DRAM and supports larger memory capacity

Performance varies by use, configuration and other factors. Configurations see appendix [5]. Results may vary

Announcing 4 Network Optimized SKU's

8351N

36 Cores
2.4GHz, 225W

First Single Socket-Optimized SKU

6338N

32 Cores
2.2GHz, 185W

5G and vRAN Optimized

6330N

28 Cores
2.2GHz, 165W

Targeted Mainstream Performance/$/W

5318N

24/20 Cores
2.1GHz, 150/135W

Intel SST-PP Enabling TDP Savings

Low Latency ▪ High Throughput ▪ Deterministic Performance

A Quick Primer on AVX-512

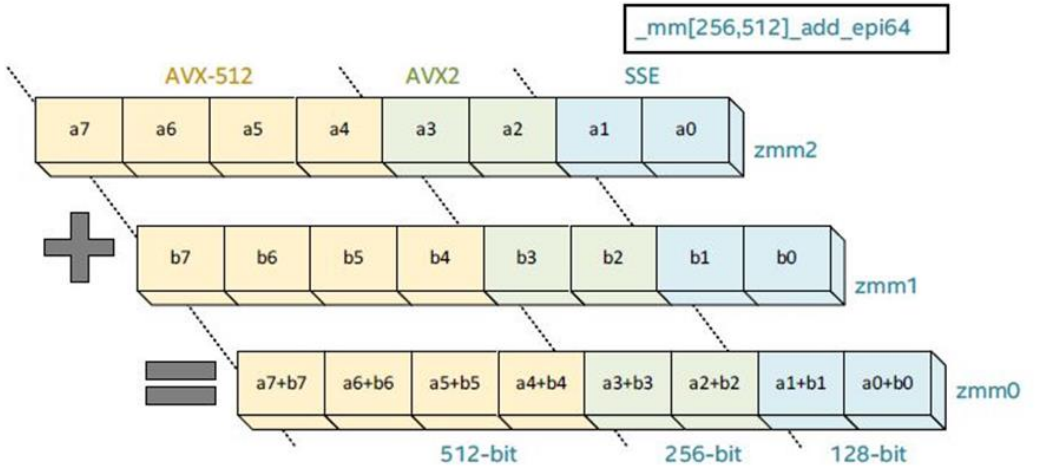
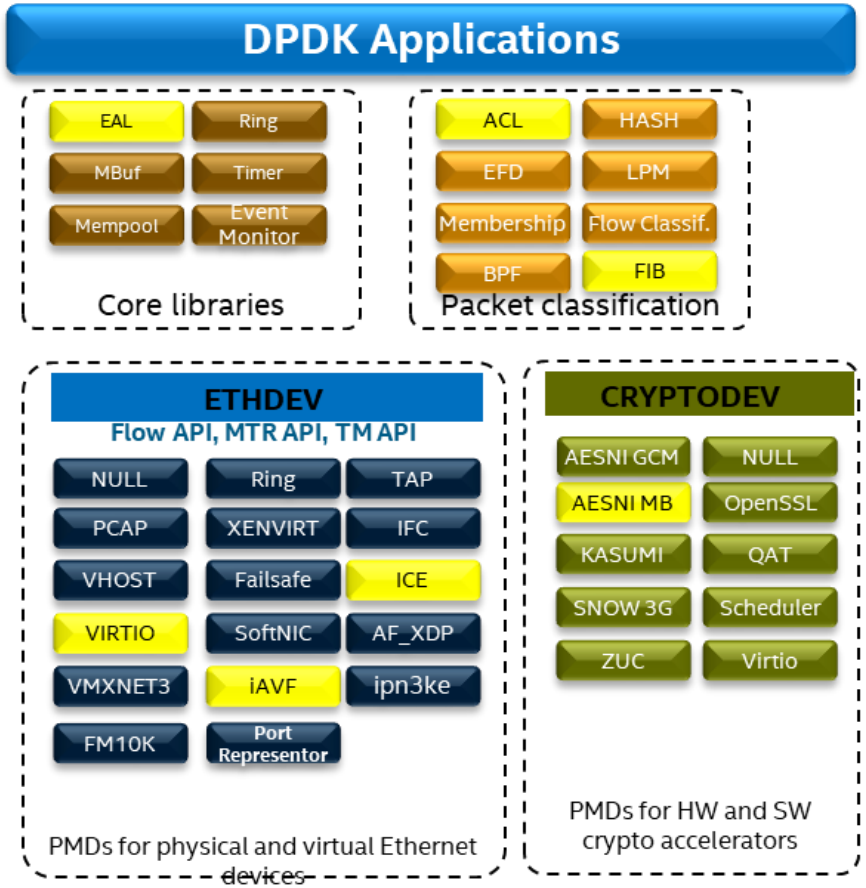


Figure 1. Vector Addition with Intel® SSE, Intel® AVX2, and Intel® AVX-512 Instruction Sets

- <u>Acceleration:</u>
- Intel® AVX-512 is a powerful SIMD instruction set
- Figure 1 shows packed 64-bit integer arithmetic doubling in the throughput with each Intel® Architecture SIMD generation, from Intel® Streaming SIMD Extensions (Intel® SSE) through to Intel® AVX-512 instructions, culminating in Intel® AVX-512 instruction's raw power to process 512 bits of data in each operation.

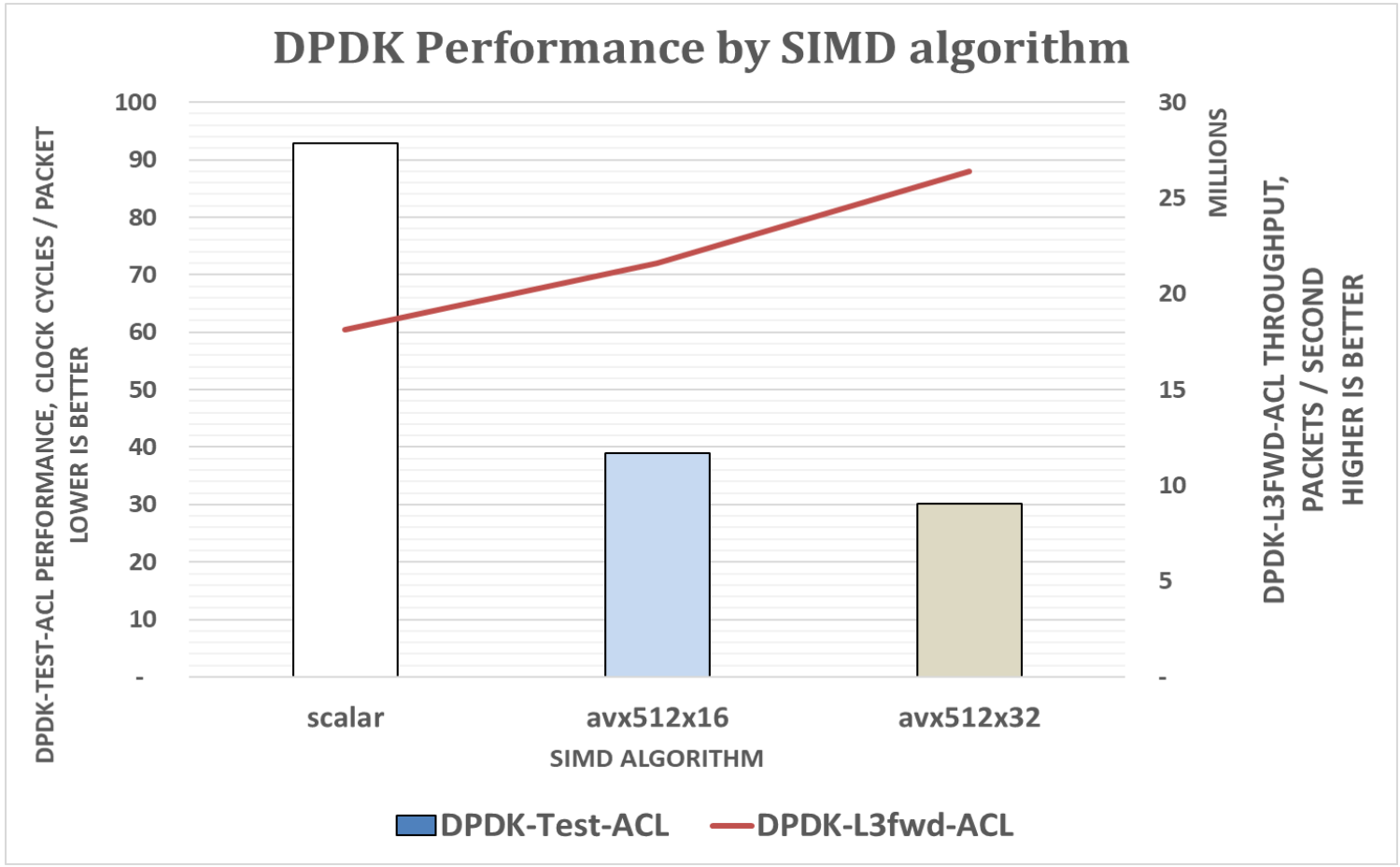
DPDK Enabling (DPDK 20.11 Release)



- New APIs have been added for applications to query and enable the use of Intel® AVX-512 instructions in DPDK.
- Intel® AVX-512 instruction set support has been added in the Access Control List (ACL) and Forwarding Information Base (FIB) libraries.
- Intel® AVX-512 instruction set is used to accelerate a number of DPDK's poll mode drivers (PMDs), including Intel® Ethernet 800 Series, Intel® Ethernet Adaptive Virtual Function (Intel® AVF), VIRTIO and Cryptodev AESNI* PMDs.

* Dependency on IPSEC MB Library v0.53

3x Performance Improvement w/AVX-512



* Single Core, Single Thread, 64-byte Packet Performance with DPDK L3FWD-ACL and DPDK-TEST-ACL example applications, 4096 Flows, 4096 ACL Rules on 3rd Generation Intel® Xeon Scalable Processors. 3x compared to scalar lookups when tested with an ACL flow lookup microbenchmark, performance is improved by up to 1.35x when used within a Layer 3 forwarding application.

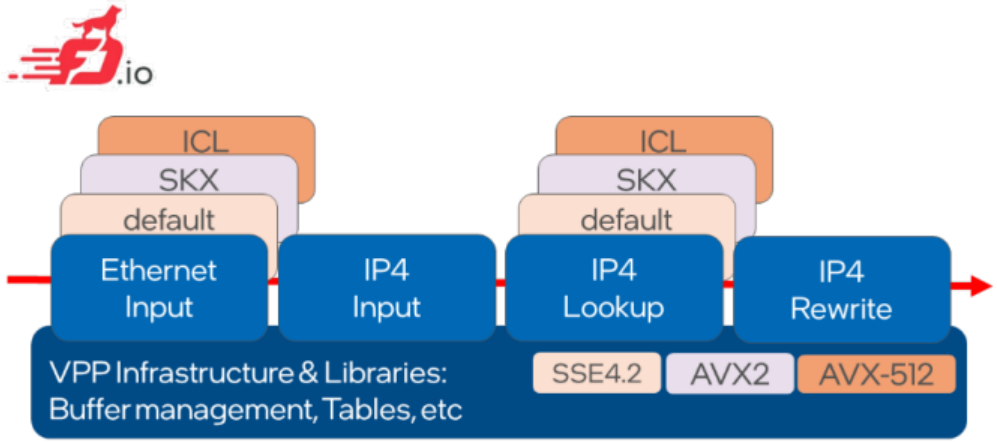
* Results have been estimated based on pre-production parts tests as of July 2021. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

FD.io Enabling

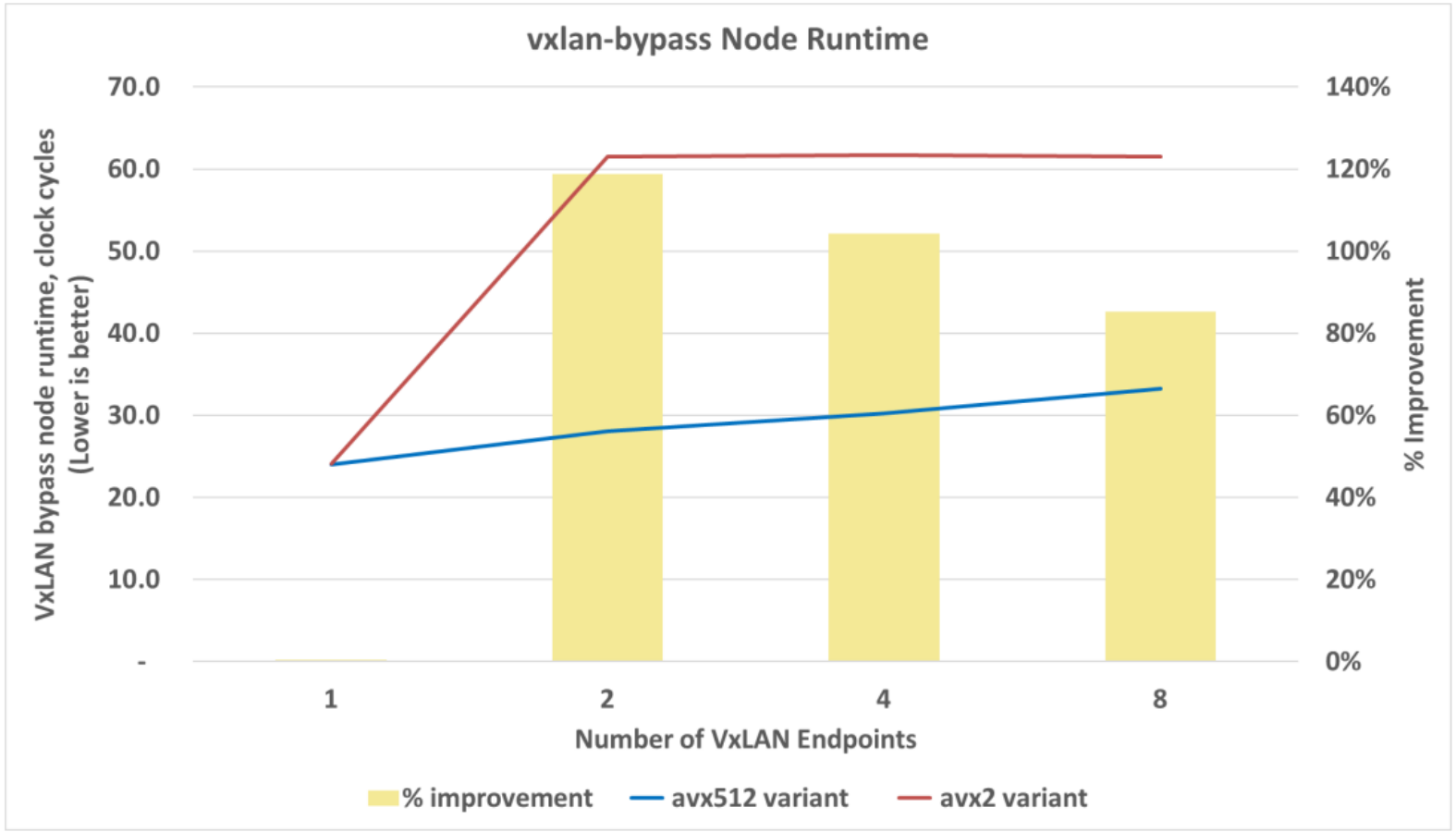
Intel® AVX-512 Support in recent FD.io Releases

- FD.io VPP support for Multi-arch Variants (see below) & IPSEC w/vAES (20.05)
- Acceleration of VPP Infra Structure with AVX-512 & IPSEC w/QAT (20.09)
- Acceleration of VXLan, Geneve & GTPU Termination & Async Crypto Worker Cores (21.01)
- Acceleration of VPP DPDK Plugin with AVX-512 (upcoming in 21.06).

- *FD.io VPP Multi-arch* variants enable architecture-specific variants of performance-sensitive graph nodes at runtime.
- Intel® AVX-512 optimizations specific are automatically enabled in supported graph-node variants.



Accelerating vTEP Termination w/AVX-512*



* Single Core, Single Thread, 64-byte Packet Performance with FD.io VPP VXLAN Tunnel Termination, up to 8 Flows, up to 8 vTEP on 3rd Generation Intel® Xeon Scalable Processors. Performance of the VxLAN decapsulation graph node improves by over 80% for 2, 4, and 8 VxLAN Endpoints. This translates into a performance improvement of up to 11% for the entire packet processing application.

* Results have been estimated based on pre-production parts tests as of July 2021. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks

AVX 512

- Intel® AVX-512 - Instruction Set for Packet Processing Technology Guide
- Intel® AVX-512 - Packet Processing with Intel® AVX-512 Instruction Set Solution Brief
- Intel® AVX-512 - Writing Packet Processing Software with Intel® AVX-512 Instruction Set Technology Guide
- https://networkbuilders.intel.com/accelerate-packet-processing-using-intel-advanced-vector-extensions-512-intel-avx-512

Compute Architecture

New instructions

Cryptography

- Big-Number Arithmetic (AVX-512 Integer IFMA)
- Vector AES and Vector Carry-less Multiply Instructions
- Galois Field New Instructions (GFNI)
- SHA-NI

Compression/Decompression and Special SIMD

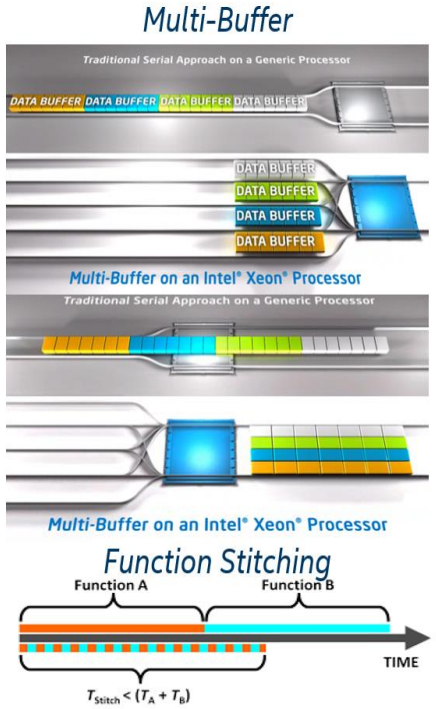
- Bit Algebra
- VBMI – Vector Bit Manipulation Instruction

New SIMD ISA Utilizing AVX512 on ICX

- Vector CLMUL
- Vector AES
- VPMADD52
- SHA Extensions
- GFNI

+

Software / Algorithms



Ice Lake vs. Cascade Lake Per Core Performance

ECDHE x25519	4.12X
RSA Sign 2048	5.63X
ECDHE p256	2.73X
AES-CTR	3.84X
AES-CMAC	3.78X
AES-XTS	3.5X
AES-GCM	3.34x
ECDSA Sign p256	1.9X
CRC	2.3X
ZUC	1.5X

Performance varies by use, configuration and other factors. Configurations see appendix [20,21,24]

Built-in Crypto Acceleration for Encryption-Intensive Workloads

Encryption made flexible

Intel Crypto Acceleration

- New instructions and architectural features parallelize execution of encryption functions
- Reduces penalty of implementing pervasive data encryption
- Higher throughput with fast and strong encryption for AVX-512 ISA
- Increases performance of encryption-intensive workloads such as SSL web server, 5G infrastructure and VPN/firewalls

4.2x

Encrypted Web Server

More TLS encrypted web server connections per second; more content delivered per server

NGINX

1.94x

Vector Packet Processing

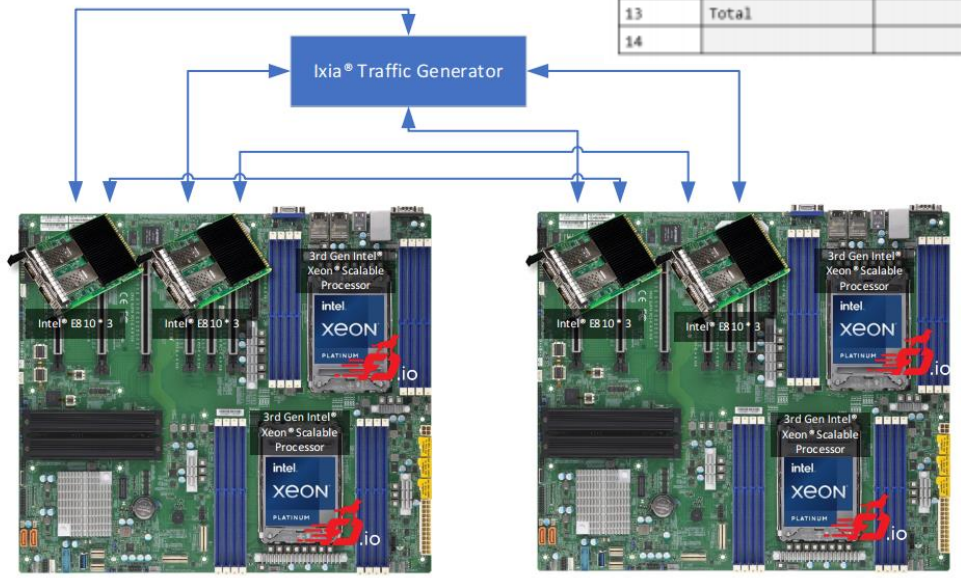
More encrypted packets processed per second; higher network and VPN capacity per node

ipsec

Performance varies by use, configuration and other factors. Configurations see appendix [17, 51]

1Tbps IPsec on Intel® Xeon® Scalable Processors

	Port Name	Line Speed	Line State	Frame Tx.	Valid Frame Rx.	Frame Delta	Loss	Frame Tx. Rate	Frame Rx. Rate	Tx. L1 Rate (bps)	Rx. L1 Rate (bps)
1	VPP Ipsec port 1	100GE	Link Up	243,937,923	243,937,063	860	0.000	10,177,203.000	10,177,188.000	84,999,999,456.000	84,999,874,176.000
2	VPP Ipsec port 2	100GE	Link Up	243,937,922	243,937,234	688	0.000	10,177,202.500	10,177,207.000	84,999,995,280.000	85,000,032,864.000
3	VPP Ipsec port 3	100GE	Link Up	243,937,922	243,937,256	666	0.000	10,177,202.500	10,177,196.500	84,999,995,280.000	84,999,945,168.000
4	VPP Ipsec port 4	100GE	Link Up	243,937,923	243,937,251	672	0.000	10,177,203.000	10,177,247.000	84,999,999,456.000	85,000,366,944.000
5	VPP Ipsec port 5	100GE	Link Up	243,937,922	243,937,044	878	0.000	10,177,202.500	10,177,188.500	84,999,995,280.000	84,999,878,352.000
6	VPP Ipsec port 6	100GE	Link Up	243,937,922	243,937,066	856	0.000	10,177,202.500	10,177,185.500	84,999,995,280.000	84,999,853,296.000
7	VPP Ipsec port 7	100GE	Link Up	243,937,923	243,937,146	777	0.000	10,177,202.500	10,177,184.000	84,999,995,280.000	84,999,840,768.000
8	VPP Ipsec port 8	100GE	Link Up	243,937,923	243,936,749	1,174	0.000	10,177,203.000	10,177,229.000	84,999,999,456.000	85,000,216,608.000
9	VPP Ipsec port 9	100GE	Link Up	243,937,923	243,936,981	942	0.000	10,177,203.000	10,177,183.500	84,999,999,456.000	84,999,836,592.000
10	VPP Ipsec port 10	100GE	Link Up	243,937,922	243,937,187	735	0.000	10,177,202.500	10,177,181.500	84,999,995,280.000	84,999,819,888.000
11	VPP Ipsec port 11	100GE	Link Up	243,937,923	243,937,080	843	0.000	10,177,203.000	10,177,189.500	84,999,995,280.000	84,999,886,704.000
12	VPP Ipsec port 12	100GE	Link Up	243,937,923	243,937,161	762	0.000	10,177,203.000	10,177,206.000	84,999,999,456.000	85,000,024,512.000
13	Total					9,853	0.000	122,126,433.000	122,126,386.000	1,019,999,968,416.000	1,019,999,575,872.000
14								Aggr. Frame Rate =	244,252,772.000	Aggr. L1 Rate (bps) =	2,039,999,151,744.000



A: Brief Deployment Diagram

- https://networkbuilders.intel.com/solutionslibrary/3rd-generation-intel-xeon-scalable-processor-achieving-1-tbps-ipsec-with-intel-advanced-vector-extensions-512-technology-guide

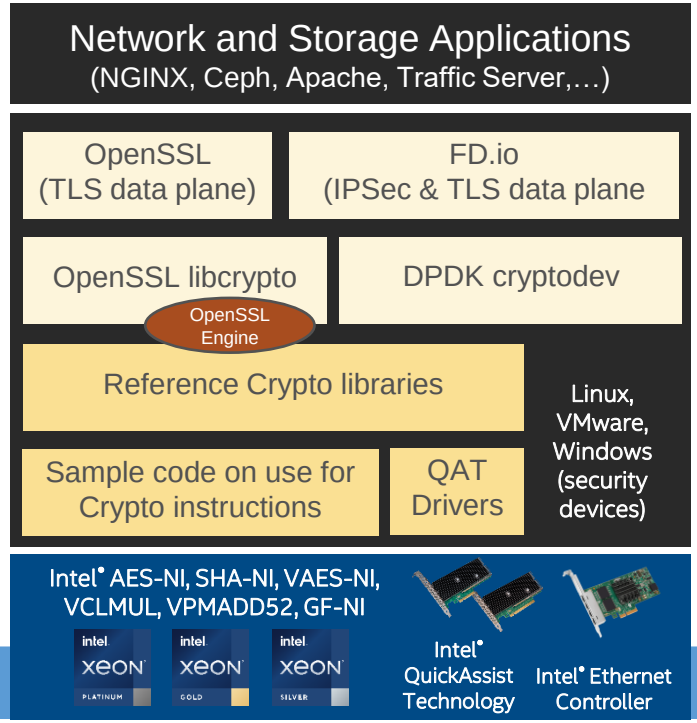
Software Optimized for Crypto New instructions

Broad set of workloads benefit from platform

- Network Security Appliances, Storage Appliances, Web Servers, Cloud Storage, Virtual Appliances, Web Application Firewall, SSL Proxy and many others

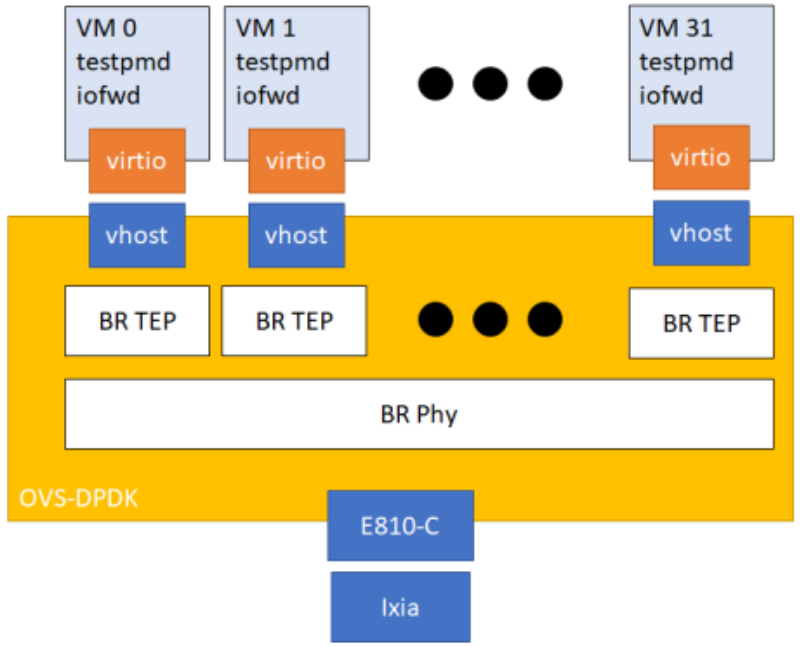
Network Transformation Experience Kit

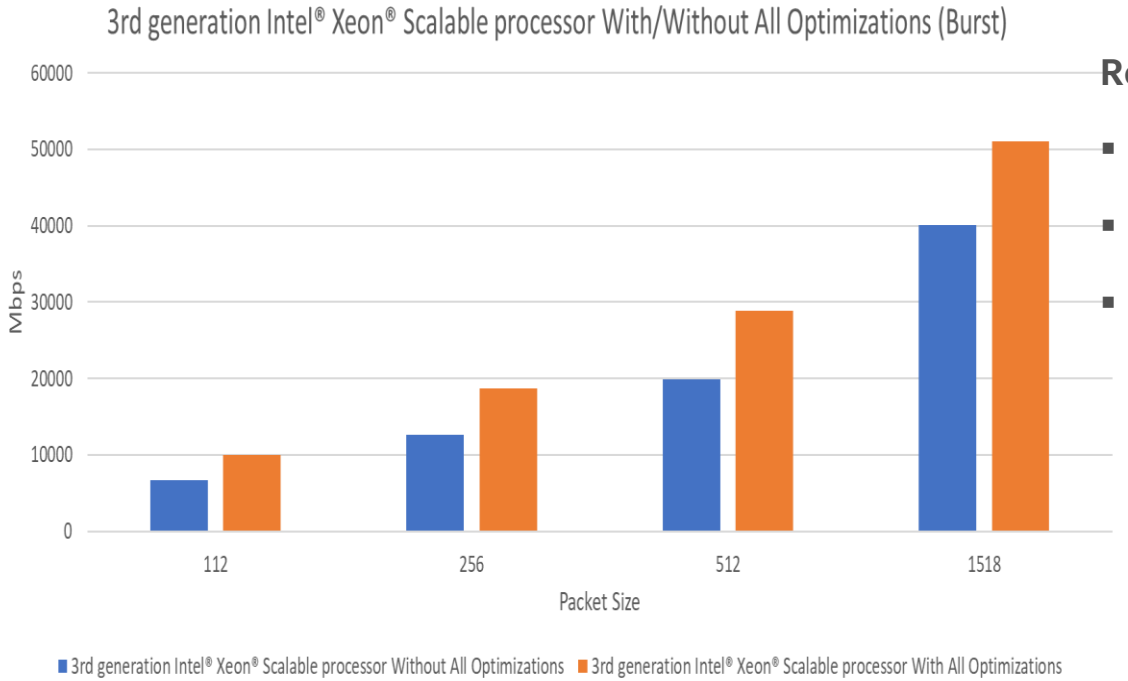
- Software reference, collateral, and benchmarks for 1TBit IPSec gateway implementation (FD.io) consuming IPSec multi-buffer library, ICX-SP PCIe Gen4 connectivity, Intel 800-Series Ethernet controllers for IPSec acceleration, ICX-SP cores with new instructions, and Intel® QuickAssist Technology



SW Resource	Description
Sample code (1/2)	ICX Crypto NI sample code for developers maintaining crypto libraries (e.g. BoringSSL, GnuTLS, WolfCrypt/WolfSSL, ...)
IPSec Multi-Buffer Lib & IPP Multi-Buffer Lib	Symmetric and Asymmetric crypto multi-buffer libraries with ICX Crypto NI optimizations including AES-GCM, RSA, X25519...
Intel OpenSSL Engine	OpenSSL Async Engine integration with crypto multi-buffer libraries with ICX Crypto NI optimizations for all applications linking with OpenSSL (virtual appliance, VNFs, ... NGFW, SSL Proxy, ..). Also on the way to RHEL to link two multi-buffer libraries in EPEL model.
Intel ISAL GitHub	Intel Storage Acceleration Library (ISA-L version 2.24) Application developers of storage workloads
DPDK	DPDK CryptoDev integration with IPSec multi-buffer libraries for AES-GCM acceleration, performance can be showed by IPSec example code.
FD.io	FD.io full IPSec dataplane solution TEMs/ISVs adding IPSec/SSL/TLS functionality to solution/product. Benchmarks for 1TBit IPSec on Whitley Platform.
Linux kernel crypto library (WIP)	Linux kernel crypto library Developers who will consume the Linux kernel crypto libraries

Open Virtual Switch





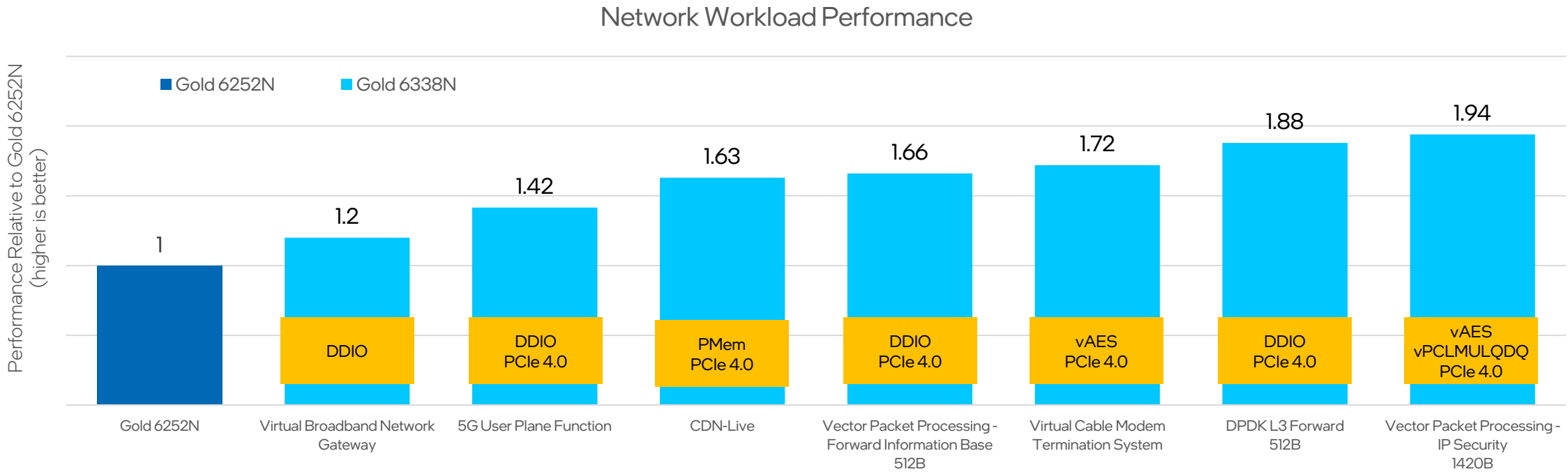
Recent Improvements

- Intel® SST-BF
- VIRTIO 1.1
- Intel® AVX-512
  - OVS DPCLS
  - DPDK ICE Vector PMD
  - OVS DPIF/MFEX *

- Intel® SST-BF - 3rd Gen Intel Xeon Scalable processor benchmarked (6338N) with a base freq of 2.2 GHz and SST-BF freq of 2.4 GHz.
- DPDK Virtio 1.1 (packed) backend in DPDK 20.11.
- Intel® AVX-512 Vector ICE PMD E800 CVL PMD in DPDK 20.11.
- Intel® AVX-512 patches are applied to the datapath interface (DPIF) and miniflow extract or packet parsing code (MFEX).
  * These patches have not yet been up-streamed, they are publicly available and under discussion with the OVS community.
- Full details are available in the paper: https://networkbuilders.intel.com/solutionslibrary/open-vswitch-optimized-deployment-benchmark-technology-guide

Network Workload Performance for Comms Infrastructure

Featuring new Intel® Xeon® Gold 6338N processor

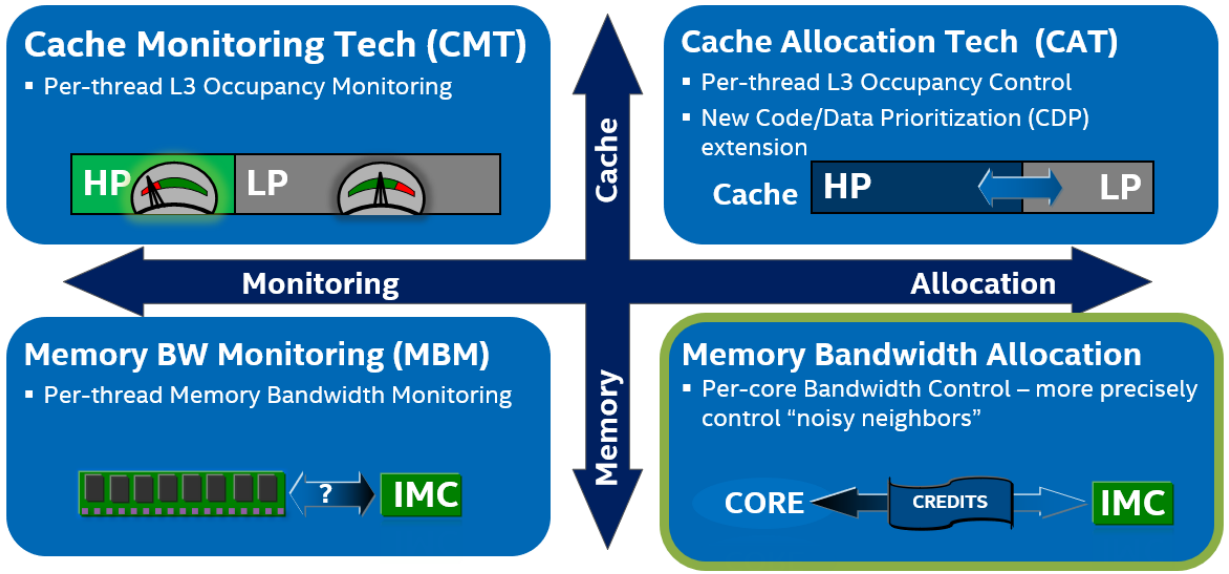


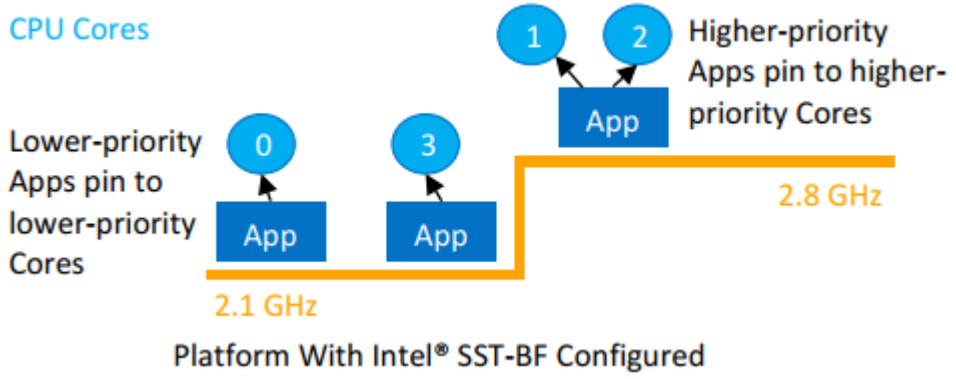
- 3rd Gen Intel® Xeon® Scalable processors, N SKUs, are engineered specifically for the needs of modern network workloads. They're targeted for low latency, high throughput, deterministic performance for our best Perf/Watt
- Ranging from 20 to 36 latest-generation cores, and 135W to 225W, these processors offer higher base frequency for greater throughput for VNF's and lower power for dense or constrained physical deployments
- 1.62x average performance gain on a range of broadly-deployed network workloads vs prior generation

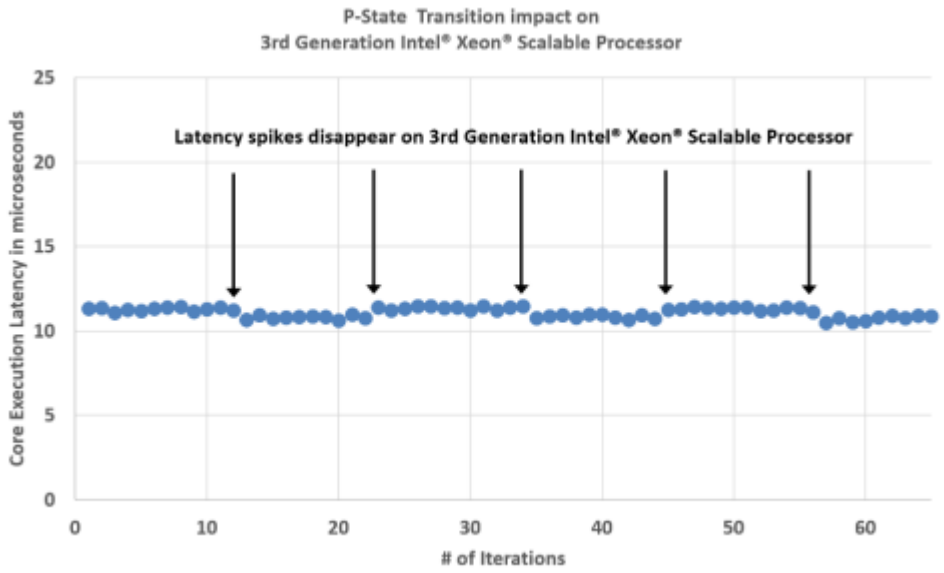
Performance varies by use, configuration and other factors. Configurations see appendix [8]

Real-Time (Deterministic) Application

- Intel RDT
- Intel SST BF
- Intel Fast Core Frequency Change

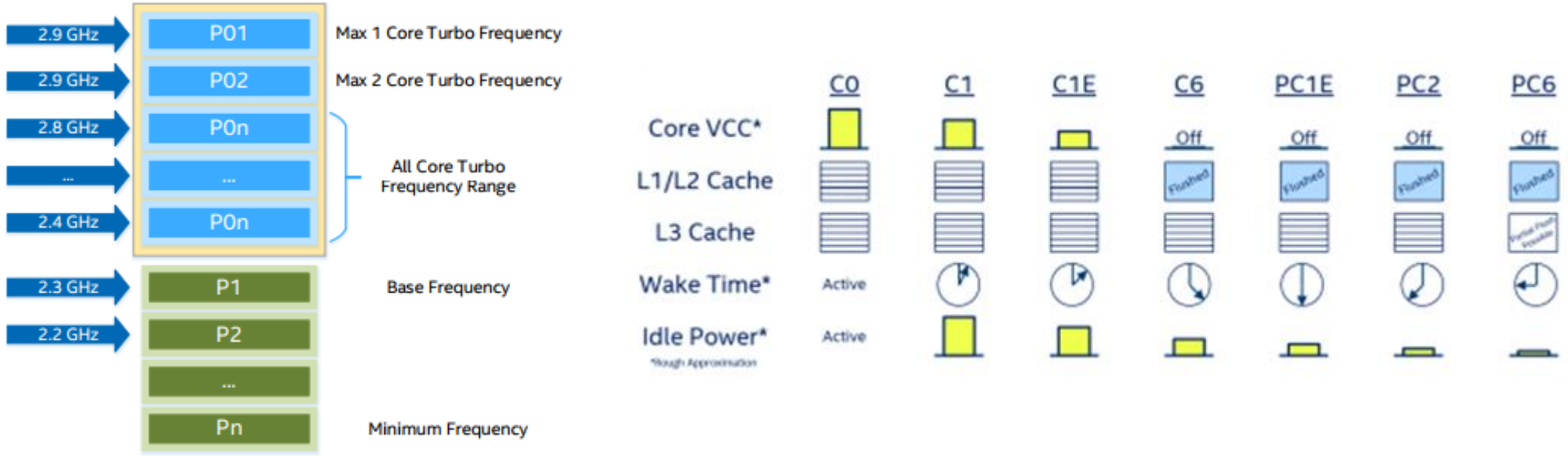






- https://networkbuilders.intel.com/solutionslibrary/power-management-enhanced-power-management-for-low-latency-workloads-technology-guide

Intel CPU Power Management



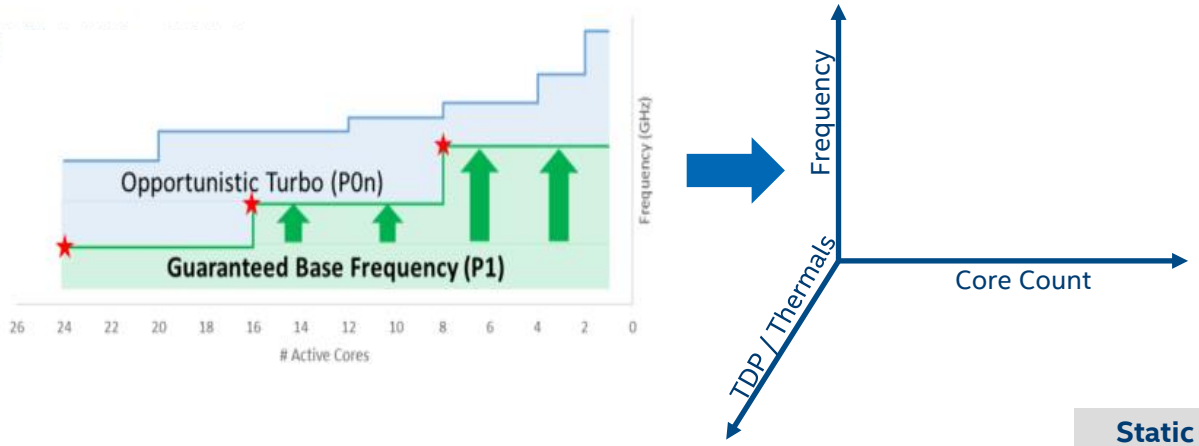
DPDK power management	**monitor**	This will use rte_power_monitor() function to enter a power-optimized state (subject to platform support).
	pause	This will use rte_power_pause() or rte_pause() to avoid busy looping when there is no traffic.
	scale	This will use frequency scaling routines available in the librte_power library.

http://doc.dpdk.org/guides/prog_guide/power_man.html
http://doc.dpdk.org/guides/sample_app_ug/l3_forward_power_man.html
https://networkbuilders.intel.com/solutionslibrary/power-management-technology-overview-technology-guide

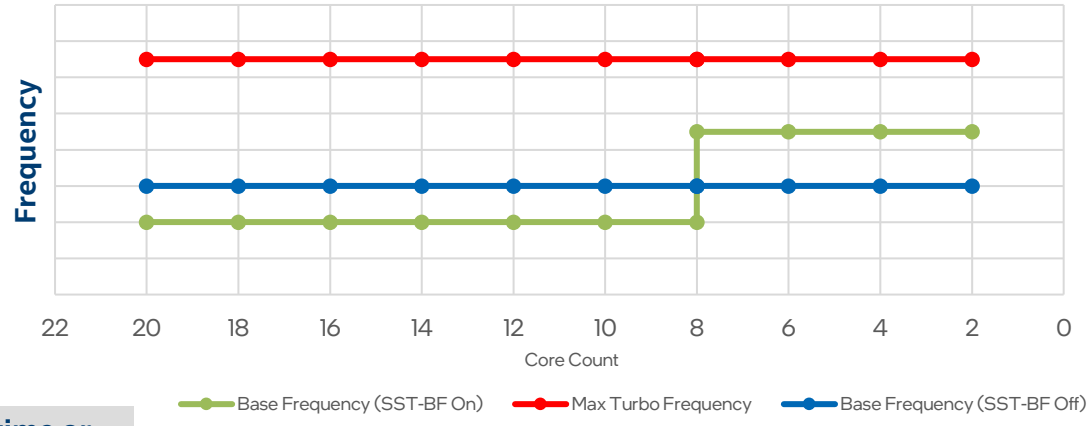
Intel® Speed Select Technology Features

Intel Speed Select Technology *is an umbrella term for a collection of features that provide more granular control over CPU performance essentially applying power/frequency where and when it is needed*

Intel® Speed Select Technology–**Performance Profile** (Intel® SST-PP)

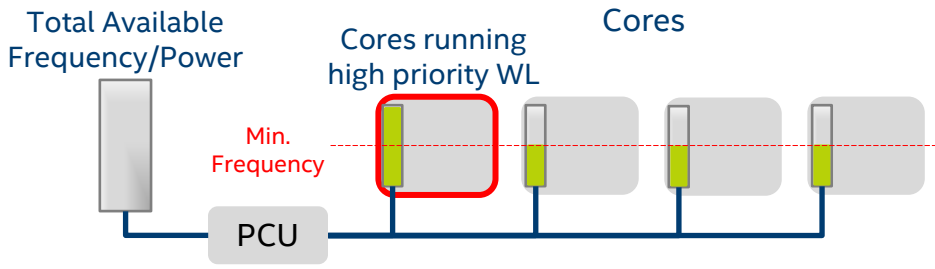


Intel® Speed Select Technology–**Base Frequency** (Intel® SST-BF)

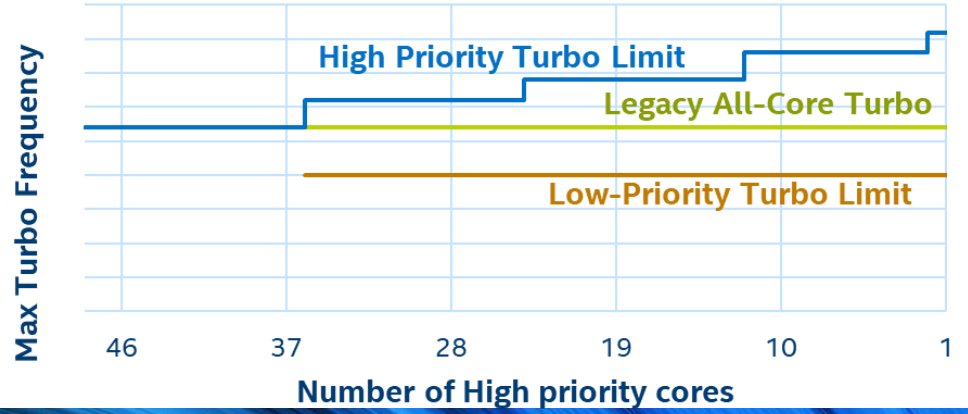


Static boot time or run time configuration

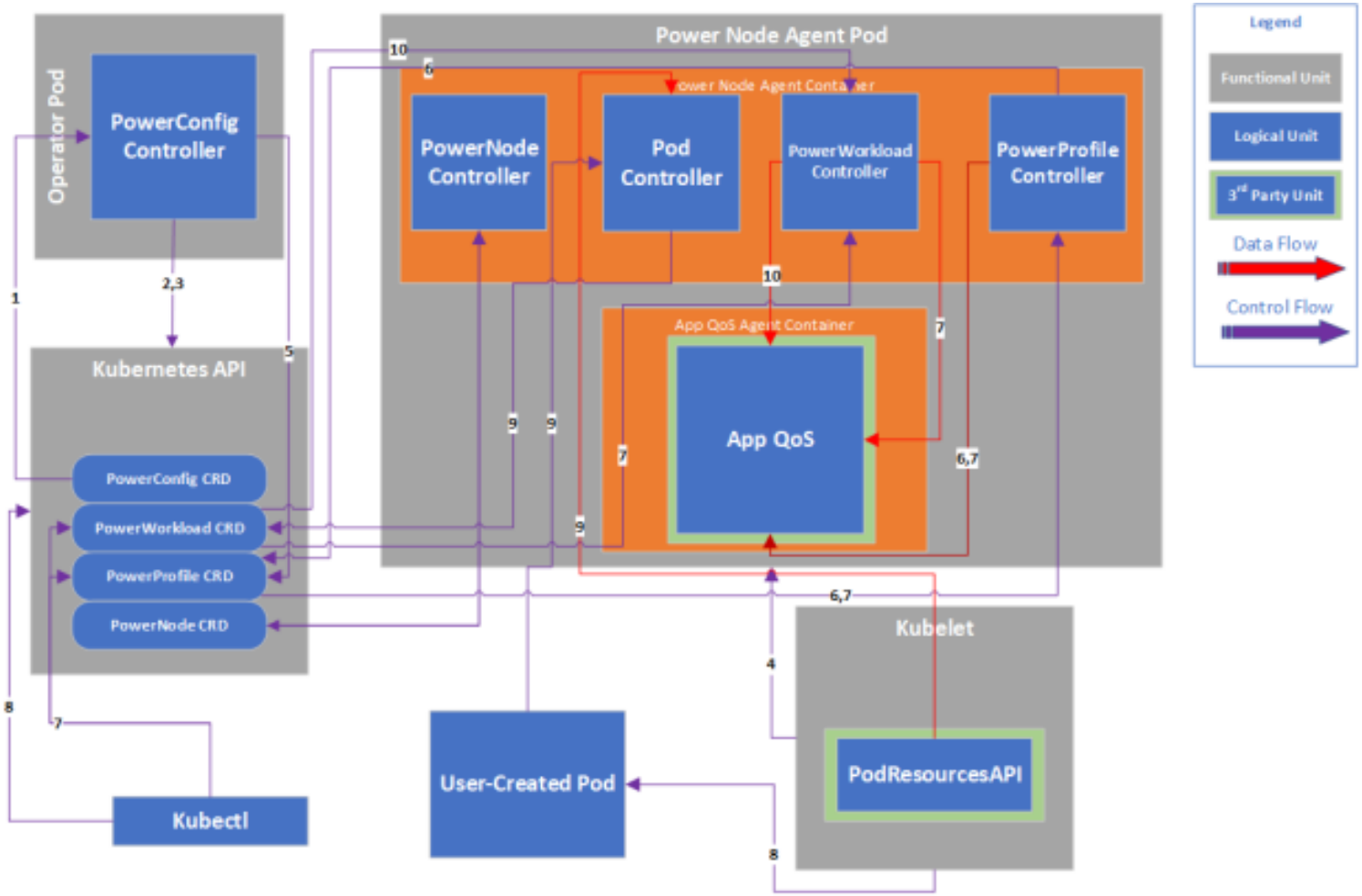
Intel® Speed Select Technology–**Core Power** (Intel® SST-CP)



Intel® Speed Select Technology–**Turbo Frequency** (Intel® SST-TF)

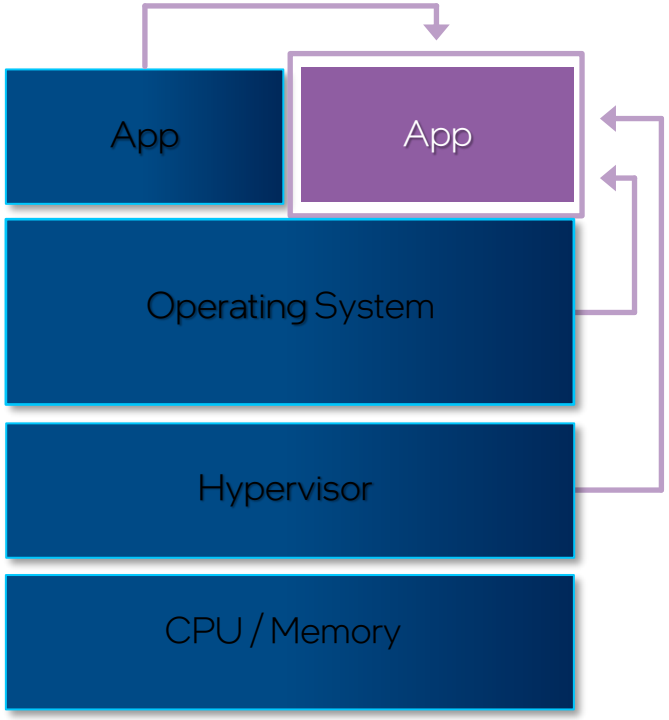


Power Manager for Kubernetes Software



https://networkbuilders.intel.com/solutionslibrary/sr-iov-with-kubernetes-networking-and-observability-technology-guide
https://github.com/intel/kubernetes-power-manager

Why Protect Data in Use?



Protect against…

Malicious insiders with escalated admin privileges

Hackers exploiting bugs in the hypervisor/OS

Third parties accessing data without owner's consent

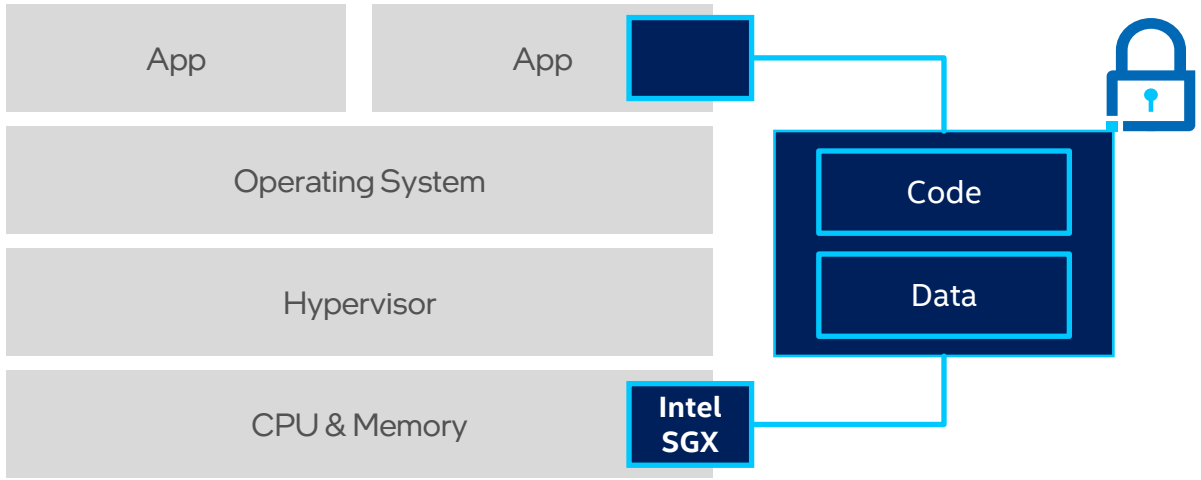
Data and Computation Exposed to…

Guest OS	Host OS	Hypervisor
Physical hardware access	Host admin	VM admin

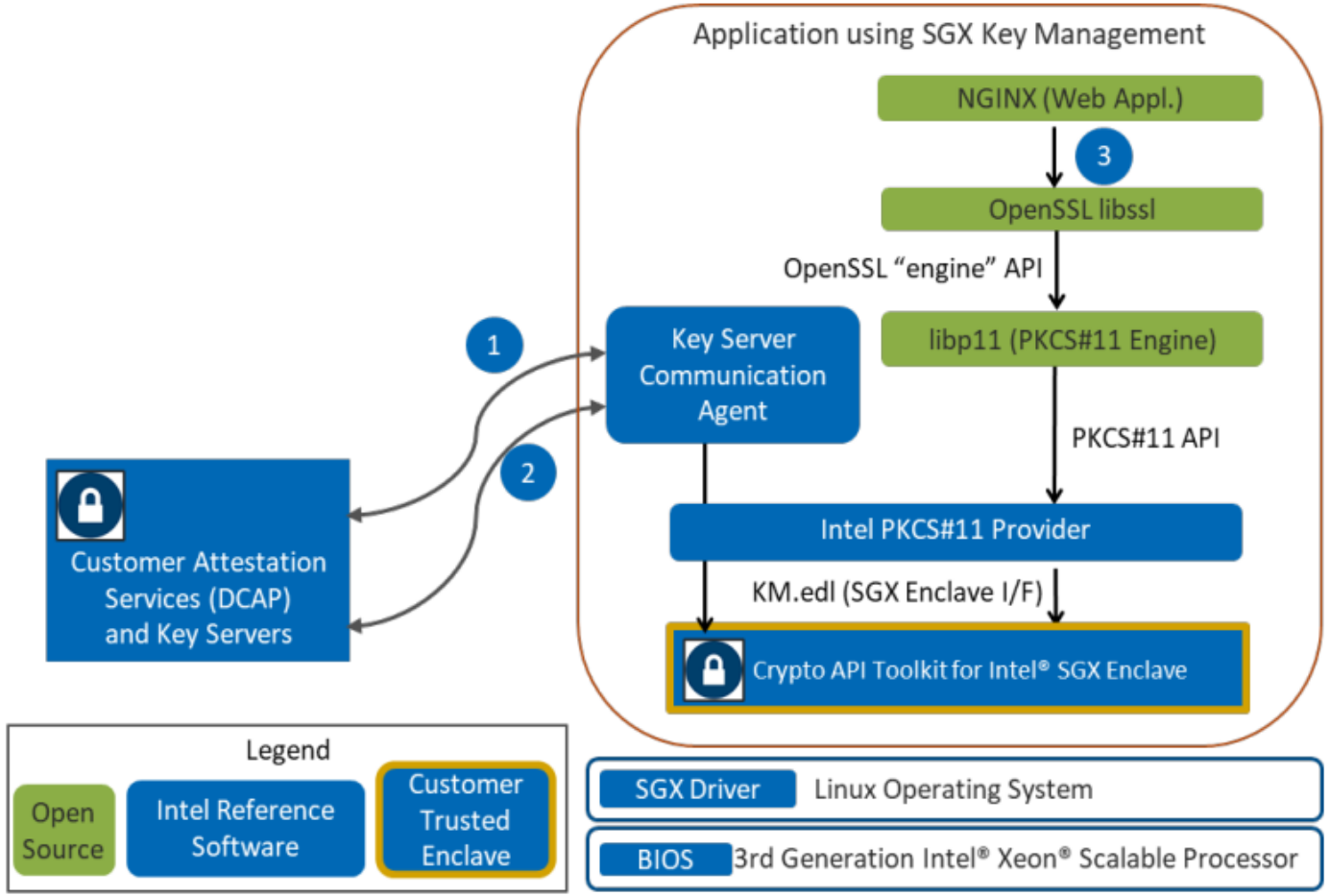
Intel Software Guard Extensions (Intel SGX)

Enables **privacy assurances for sensitive data segments** without compromising performance

Huge enclaves now support demands of mainstream workloads (up to 1TB memory spaces)



NGINX Private Key on 3rd Generation Intel SGX



https://networkbuilders.intel.com/solutionslibrary/intel-software-guard-extensions-intel-sgx-nginx-private-key-on-3rd-generation-intel-xeon-scalable-processor-user-guide

Intel Network AI Offering

DISCOVERY	Data	Develop	Deploy
of possibilities & next steps	setup, ingestion & cleaning	models using analytics/AI	into production & iterate

Network AI

Simplify Network AI Deployment with Domain Expert Support (Use-case Ref.)

1. Design Use-case solution
2. Build AI models per use case
3. Optimize AI performance
4. E2E NW & AI Solution Deployment

One Intel AI foundation

Optimized Libraries, Frameworks, Tools

- Developer Tools: ANALYTICS ZOO, OpenVINO, oneAPI, Intel Neural Compressor, TADK
- Standard Frameworks: python, TensorFlow, PyTorch, Spark ML
- Optimized Libraries:
  - Intel® Data Analytics Acceleration Library (oneDAL)
  - Intel® Deep Neural Network Library (oneDNN)
  - Intel® OneAPI Collective Library

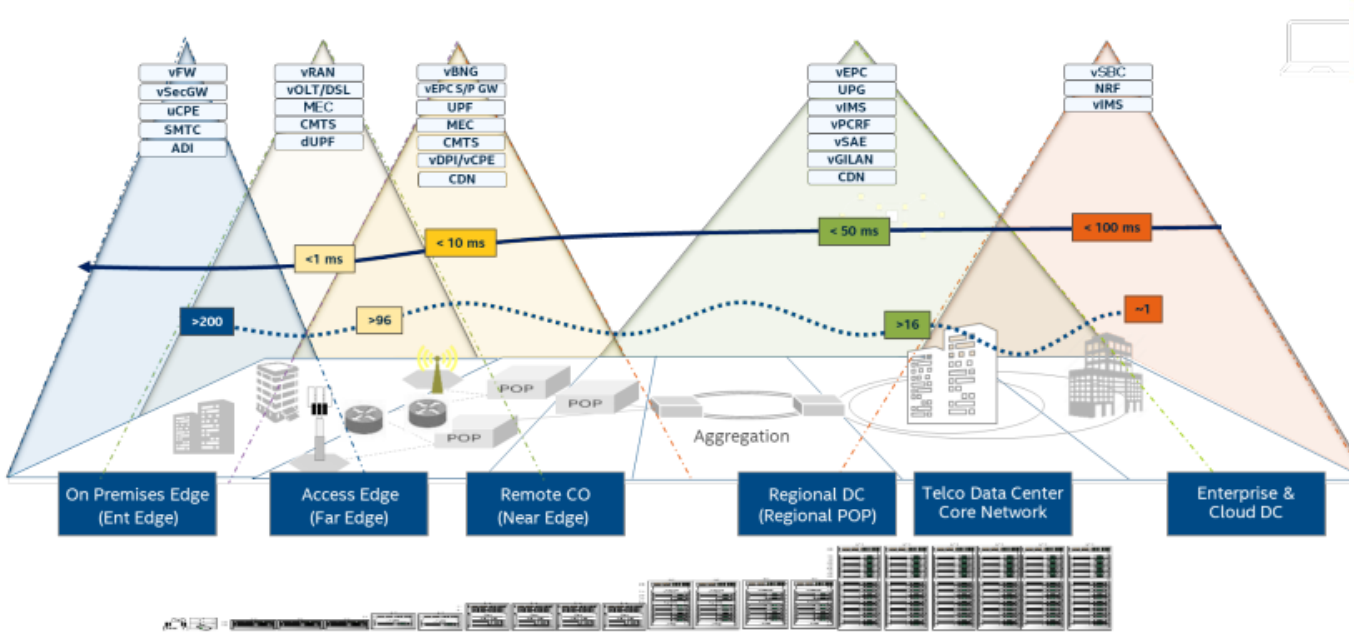
End-to-End AI Portfolio Roadmap

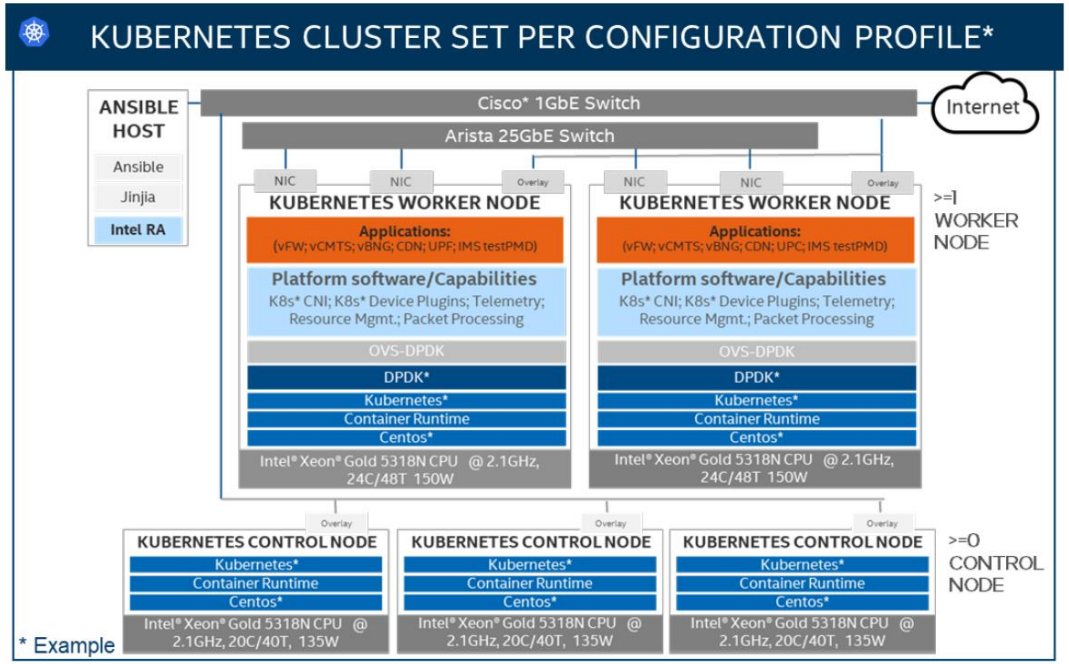
Max Optimization on Intel® Processors: intel. XEON PLATINUM, intel. ATOM + intel. AGILEX, habana, Xe

Accelerate with Purpose

Intel® DL Boost (VNNI, BF16), Intel® SSE4.1, AVX-512, AMX/TMUL (SPR)

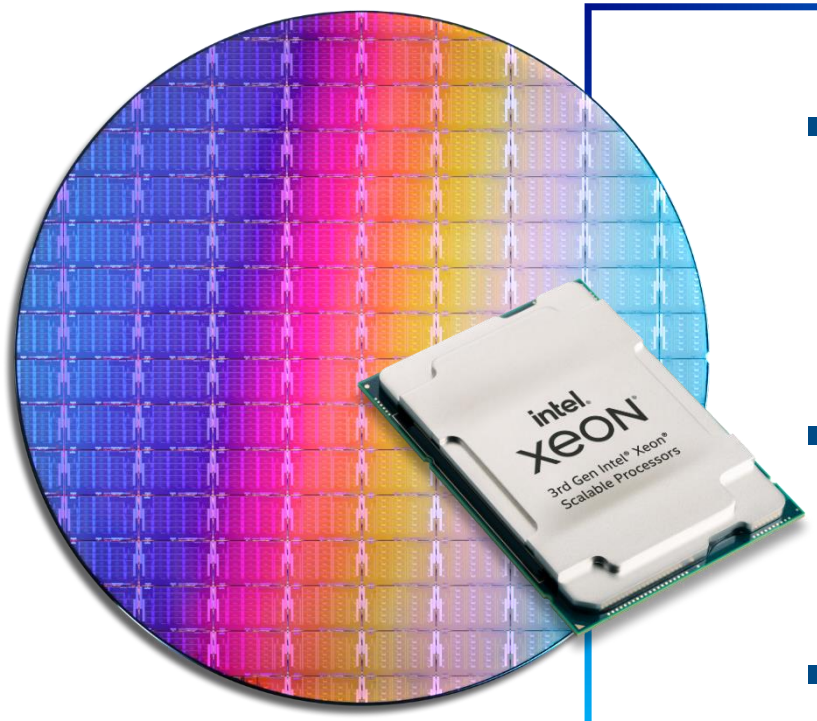
Container Bare Metal Reference Architecture





https://networkbuilders.intel.com/solutionslibrary/container-bare-metal-for-2nd-3rd-generation-intel-xeon-scalable-processor
https://github.com/intel/container-experience-kits/

Summary



- Intel's highest performing data center processor with built-in security and AI and crypto acceleration
- Unmatched portfolio of hardware and software solutions to move, store and process data
- Broadest ecosystem and decades of experience to ease customer deployments


intel®

Intel® Xeon® Performance Tuning and Solutions

- Links to tuning recipes for Nginx, Deep Learning, Video Transcode, MongoDB, Redis, HPC, database, Spark, KVM, Java…
- https://www.intel.com/content/www/us/en/developer/articles/guide/xeon-performance-tuning-and-solution-guides.html