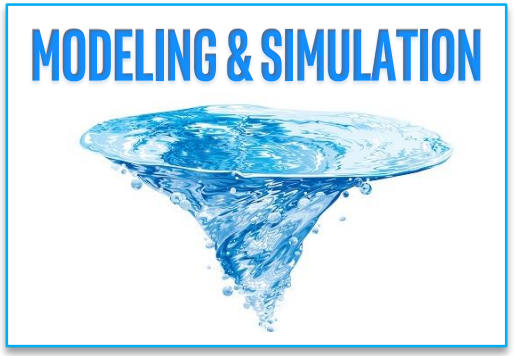
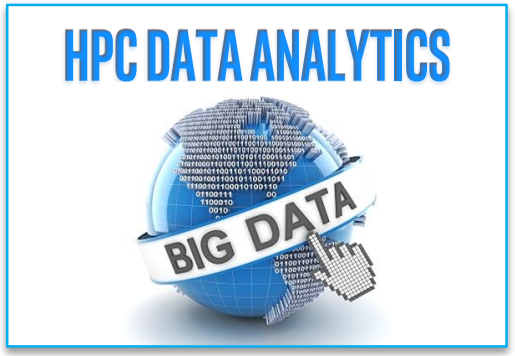
OPTIMAL HPC SOLUTIONS WITH INTEL

Nikolay Mester, HPC and CSP verticals, Eastern Europe, Intel

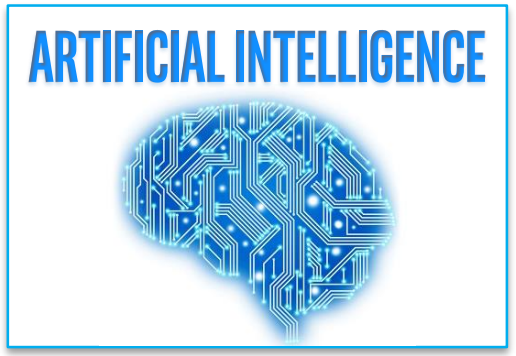
The HPC Opportunity



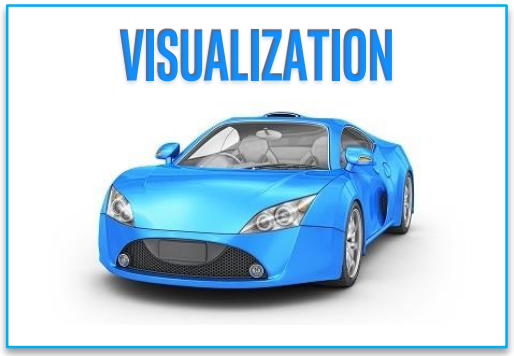
$515 average return per $1 of HPC investment[1]



18% revenue CAGR; >$3 billion in 2020[2]



55% revenue CAGR; >$47 billion in 2020[3]



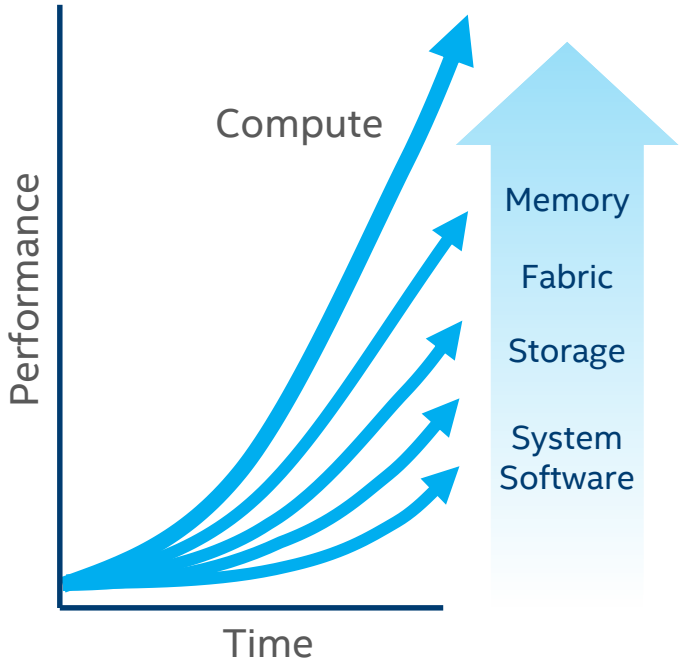
30% revenue CAGR; >$1.6 billion in 2020[4]

1 Source: Source: IDC HPC and ROI Study Update, September 2015
2 Source IDC Worldside High-Performance Data Analytics Forecast 2016-2020, June 2016
3 Source: IDC Worldwide Semiannual Cognitive/Artificial Intelligence Systems Spending Guide, Oct 2016
4 Source: MarketsandMarkets Visualization and 3D Rendering Software Market by Application, March 2016

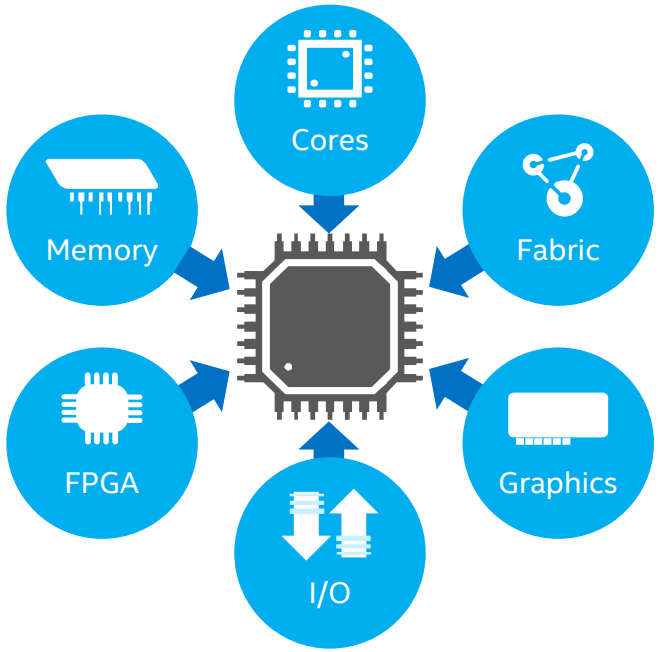
A Holistic Architectural Approach

System

Innovative Technologies

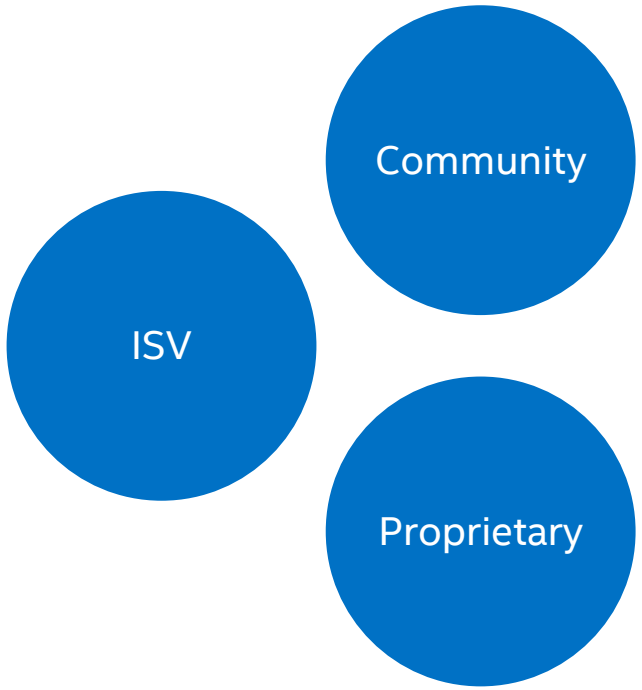


Tighter Integration

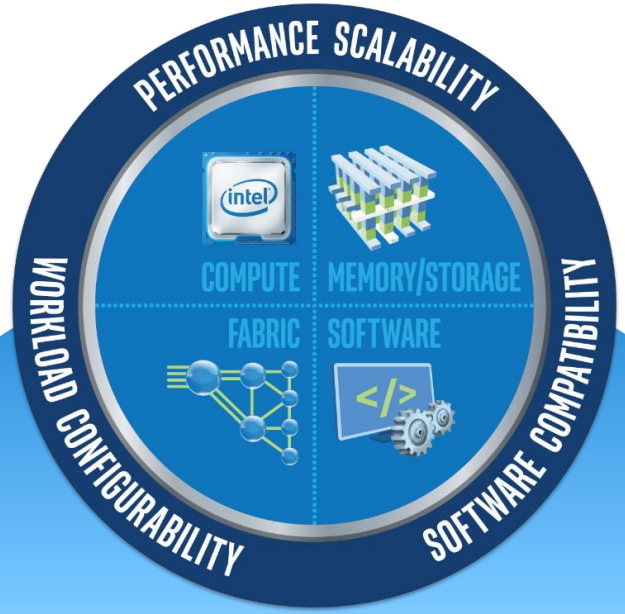


Application

Modernized Code



Key Elements of Intel® SSF

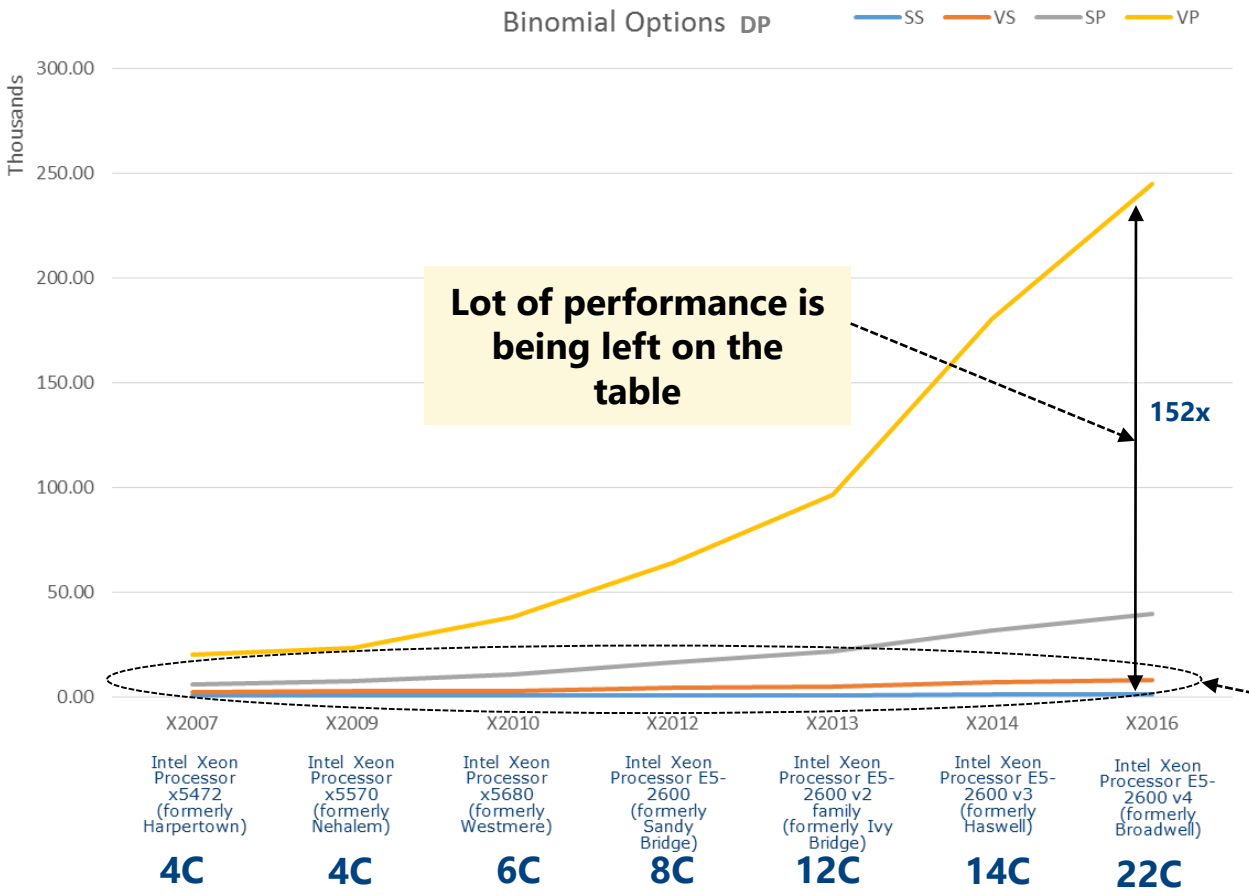


INTEL® SCALABLE SYSTEM FRAMEWORK

intel XEON PLATINUM inside™	intel XEON PHI™ inside™	INTEL® OMNI-PATH ARCHITECTURE	INTEL® HPC ORCHESTRATOR	intel SSD inside™	intel ENTERPRISE EDITION FOR* LUSTRE
MARKET LEADING[1]	HIGHLY PARALLEL	COST ADVANTAGE	INTEL SUPPORTED	FLEXIBILITY & STABILITY	EXTREME SCALABILITY

* Other names and brands may be claimed as the property of others. [1]*Source: Intel estimates.*

Code Modernization for Higher Performance



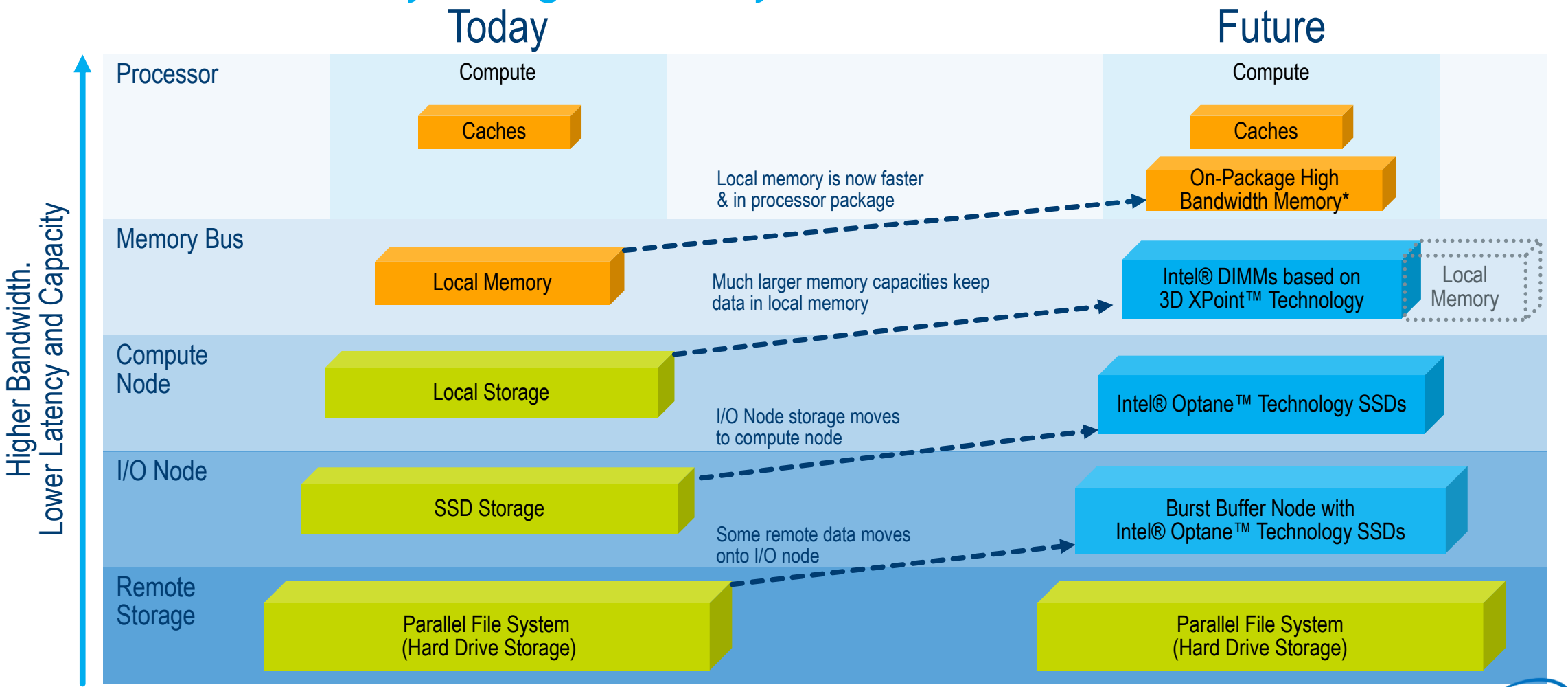
VP = Vectorized & Parallelized (MT)
SP = Scalar & Parallelized (MT)
VS = Vectorized & Single-Threaded (ST)
SS = Scalar & Single-Threaded (ST)

We believe most codes are here

Modernization (i.e. parallelization and vectorization) of your code is the solution

Tighter System-Level Integration

Innovative Memory-Storage Hierarchy



What to use for your situation?

Why Xeon Phi™?

Improve Performance

-AND/OR-

Improve ROI

Unlock Potential

Which Apps?[1]

- Scalable to >60 cores

AND

- Heavily Vectorized

-OR-

- Local memory BW bound

If yes...

Optimized for Highly-Parallel Applications

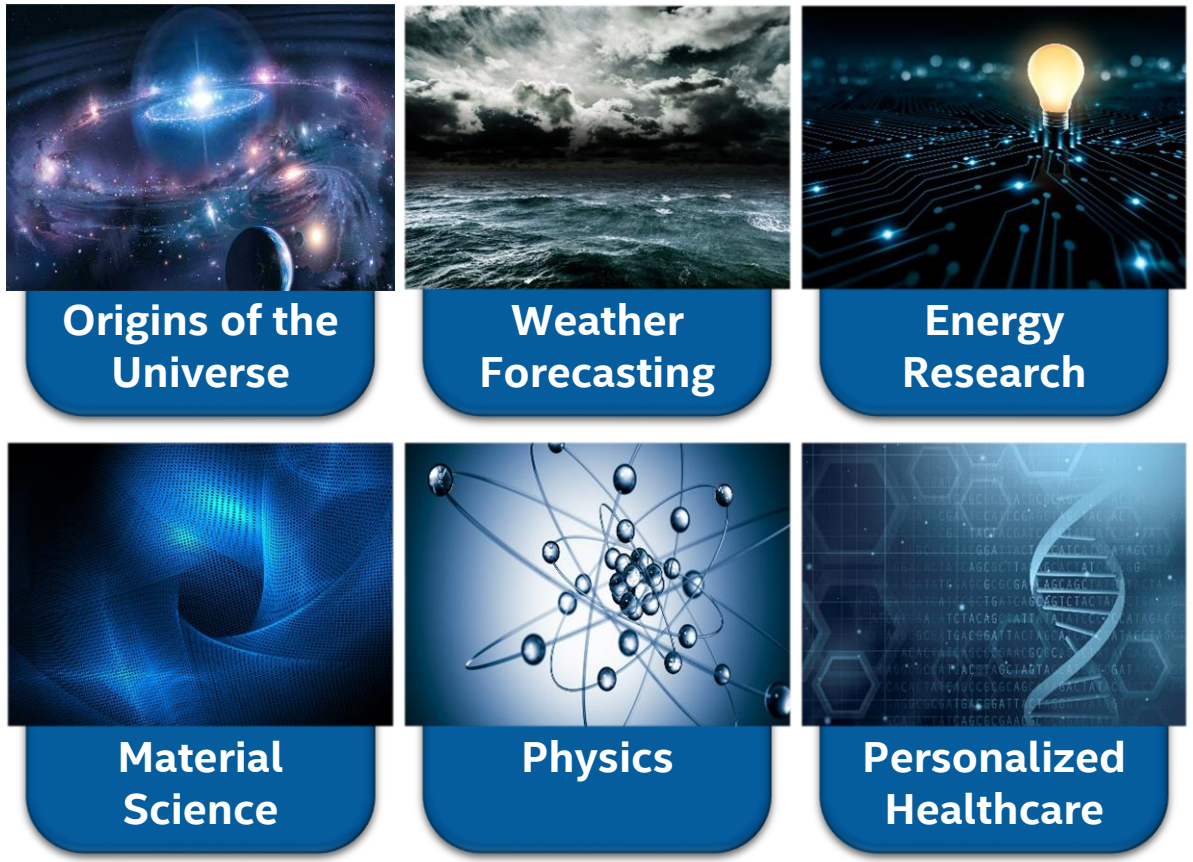
If no...

Commonly-Used Parallel Processor*

Intel® Xeon Phi™ is optimal for applications that scale to >60 cores and are highly threaded or memory bandwidth bound

[1]Performance results on Intel® Xeon Phi™ will vary depending on app characteristics. For more information, see: https://software.intel.com/sites/default/files/article/383067/is-xeon-phi-right-for-me.pdf

Intel® Xeon® Scalable Processor Enables Amazing Discoveries through HPC



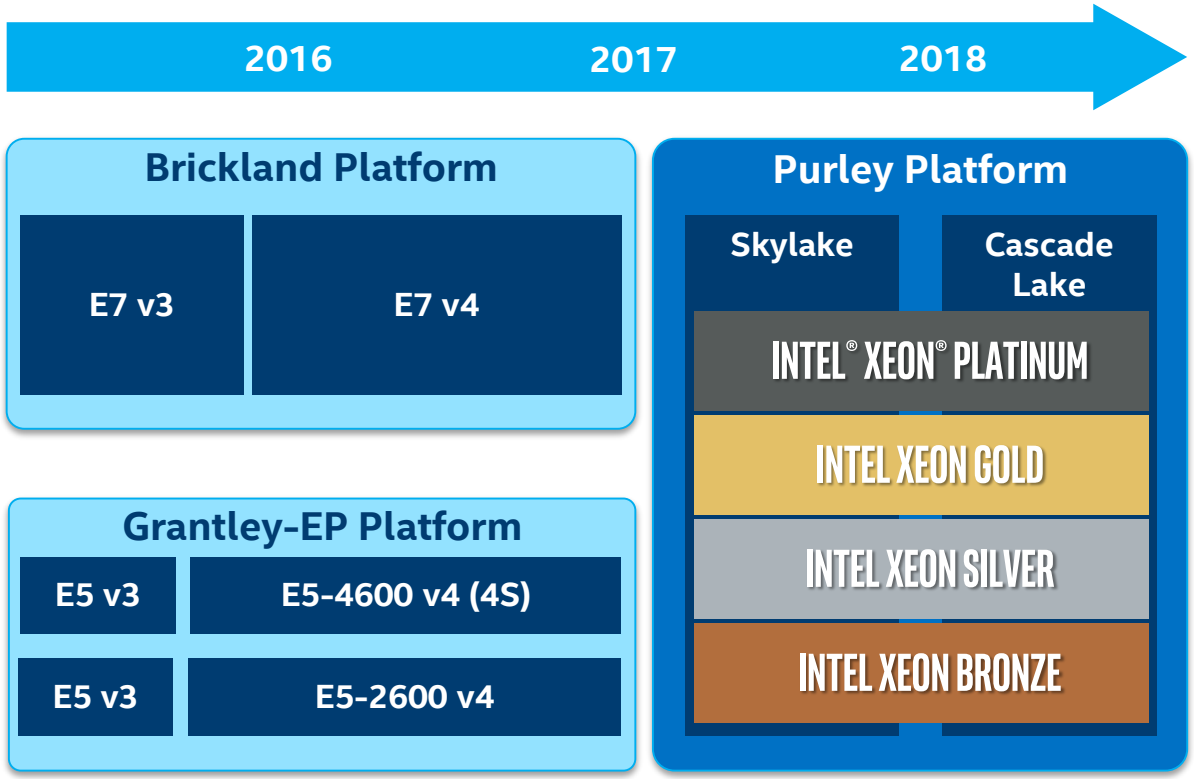
Intel® Xeon® Processor Roadmap

Intel® Xeon® Processor E7

Targeted at **mission critical** applications that value a **scale-up** system with leadership **memory capacity** and **advanced RAS**

Intel® Xeon® Processor E5

Targeted at a wide variety of applications that value a **balanced system** with **leadership** performance/watt/$



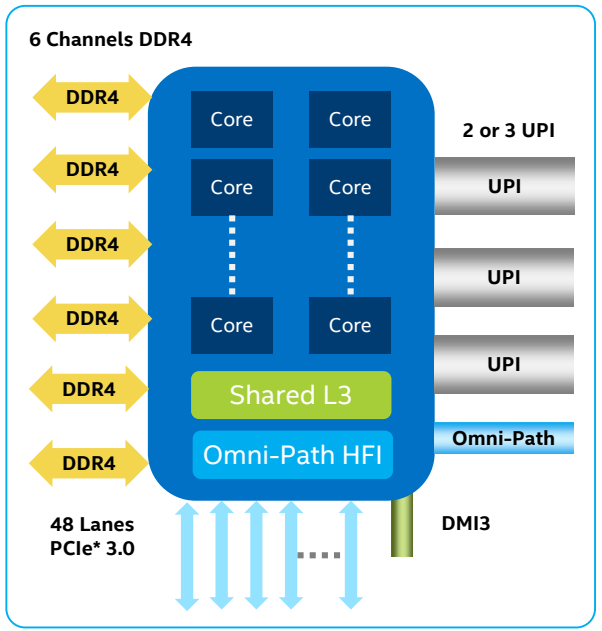
CONVERGED PLATFORM WITH INNOVATIVE SKYLAKE-SP MICROARCHITECTURE

Intel® Xeon® Scalable Processor

Re-architected from the Ground Up

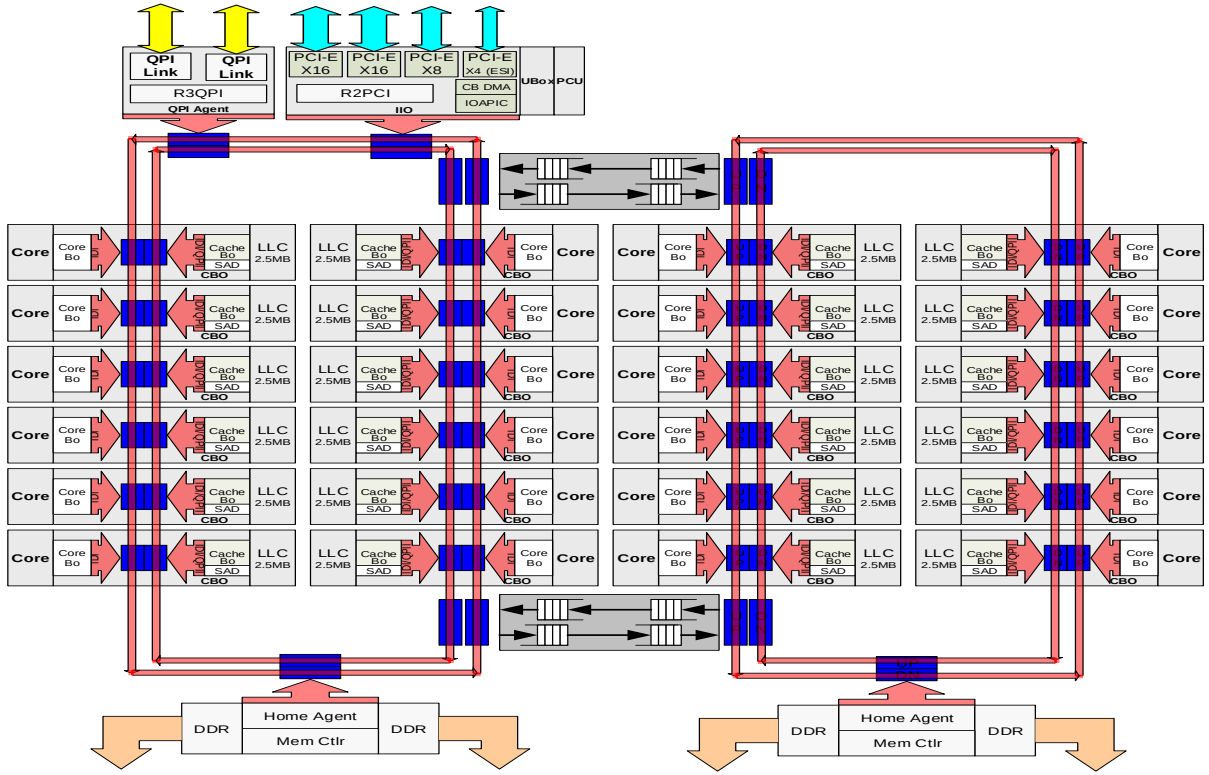
- Skylake core microarchitecture, with data center specific enhancements
- Intel® AVX-512 with 32 DP flops per core
- Data center optimized cache hierarchy – 1MB L2 per core, non-inclusive L3
- New mesh interconnect architecture
- Enhanced memory subsystem
- Modular IO with integrated devices
- New Intel® Ultra Path Interconnect (Intel® UPI)
- Intel® Speed Shift Technology
- Security & Virtualization enhancements (MBE, PPK, MPX)
- Optional Integrated Intel® Omni-Path Fabric (Intel® OPA)

Features	Intel® Xeon® Processor E5-2600 v4	Intel® Xeon® Scalable Processor
Cores Per Socket	Up to 22	**Up to 28**
Threads Per Socket	Up to 44 threads	**Up to 56 threads**
Last-level Cache (LLC)	Up to 55 MB	**Up to 38.5 MB (non-inclusive)**
QPI/UPI Speed (GT/s)	2x QPI channels @ 9.6 GT/s	**Up to 3x UPI @ 10.4 GT/s**
PCIe* Lanes/ Controllers/Speed(GT/s)	40 / 10 / PCIe* 3.0 (2.5, 5, 8 GT/s)	**48 / 12 / PCIe 3.0 (2.5, 5, 8 GT/s)**
Memory Population	4 channels of up to 3 RDIMMs, LRDIMMs, or 3DS LRDIMMs	**6 channels of up to 2 RDIMMs, LRDIMMs, or 3DS LRDIMMs**
Max Memory Speed	Up to 2400	**Up to 2666**
TDP (W)	55W-145W	**70W-205W**

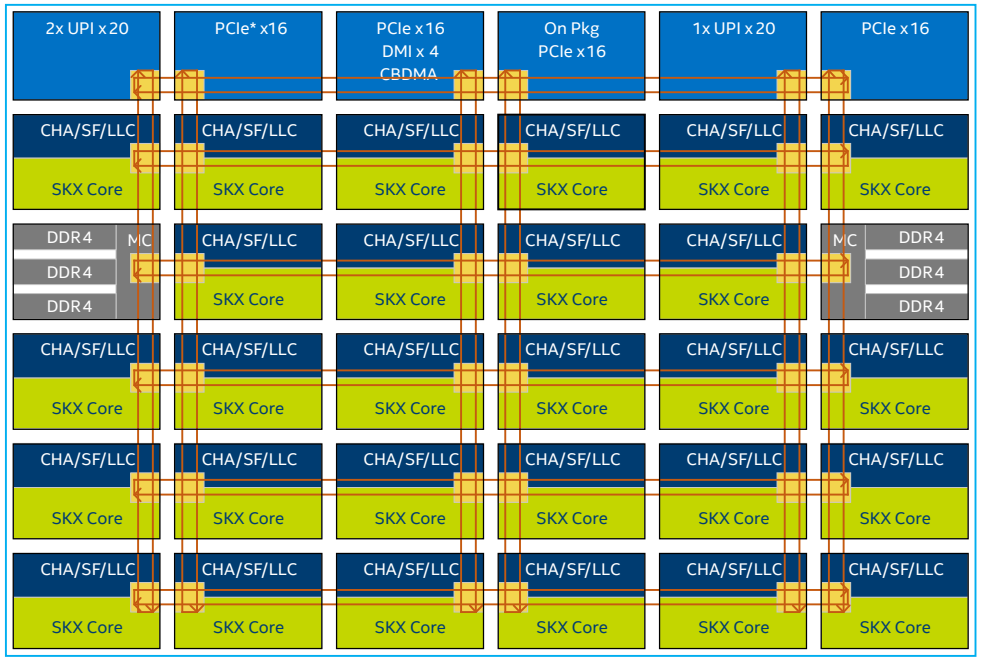


New Mesh Interconnect Architecture

Broadwell EX 24-core die



Skylake-SP 28-core die

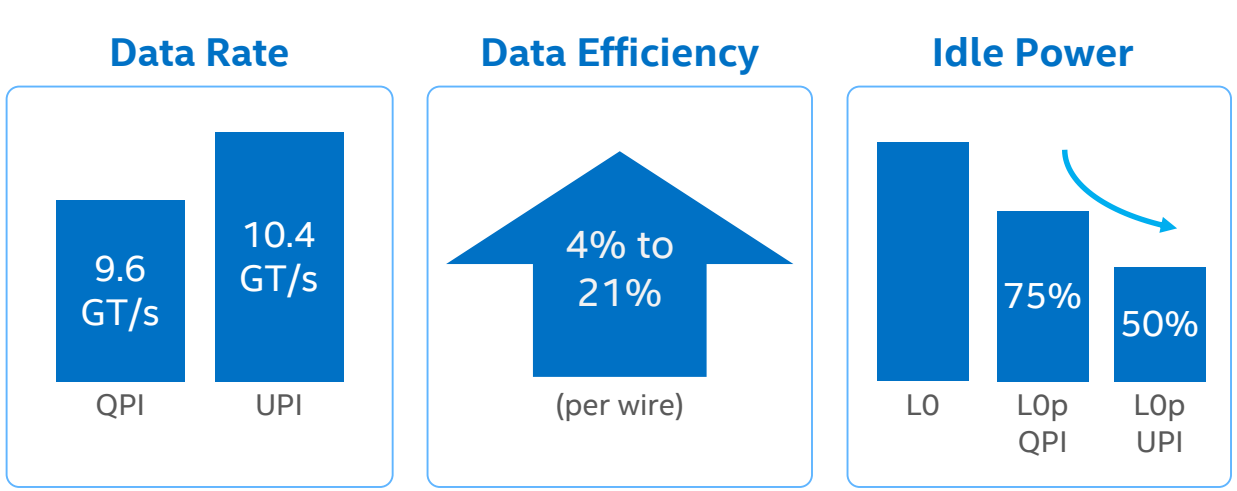


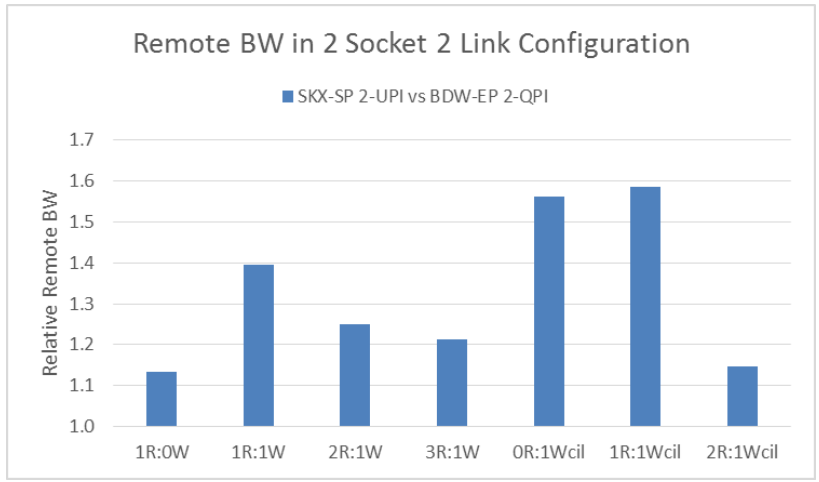
CHA – Caching and Home Agent ; SF – Snoop Filter; LLC – Last Level Cache ;
SKX Core – Skylake Server Core ; UPI – Intel® UltraPath Interconnect

MESH IMPROVES SCALABILITY WITH HIGHER BANDWIDTH AND REDUCED LATENCIES

Intel® Ultra Path Interconnect (Intel® UPI)

- Intel® Ultra Path Interconnect (Intel® UPI), replacing Intel® QPI
- Faster link with improved bandwidth for a balanced system design
  - Improved messaging efficiency per packet
- 3 UPI option for 2 socket – additional inter-socket bandwidth for non-NUMA optimized use-cases



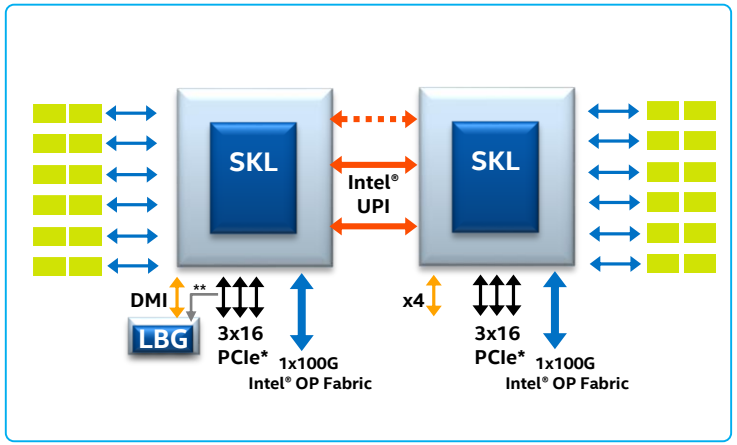


INTEL® UPI ENABLES SYSTEM SCALABILITY WITH HIGHER INTER-SOCKET BANDWIDTH

Source as of June 2017: Intel internal measurements on platform with Xeon Platinum 8180, Turbo enabled, UPI=10.4, 6x32GB DDR4-2666, 1 DPC, and platform with E5-2699 v4, Turbo enabled, 4x32GB DDR4-2400, RHEL 7.0. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.
Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

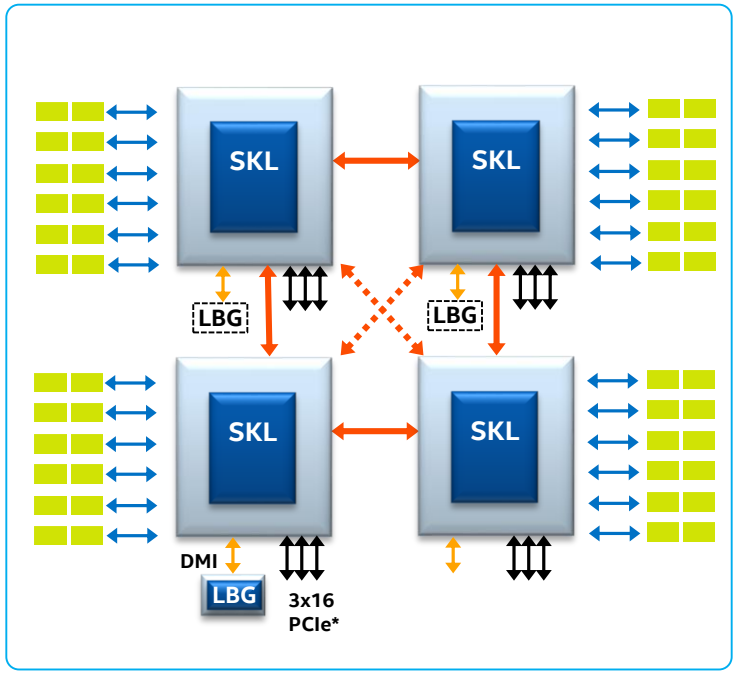
Platform Topologies

2S Configurations



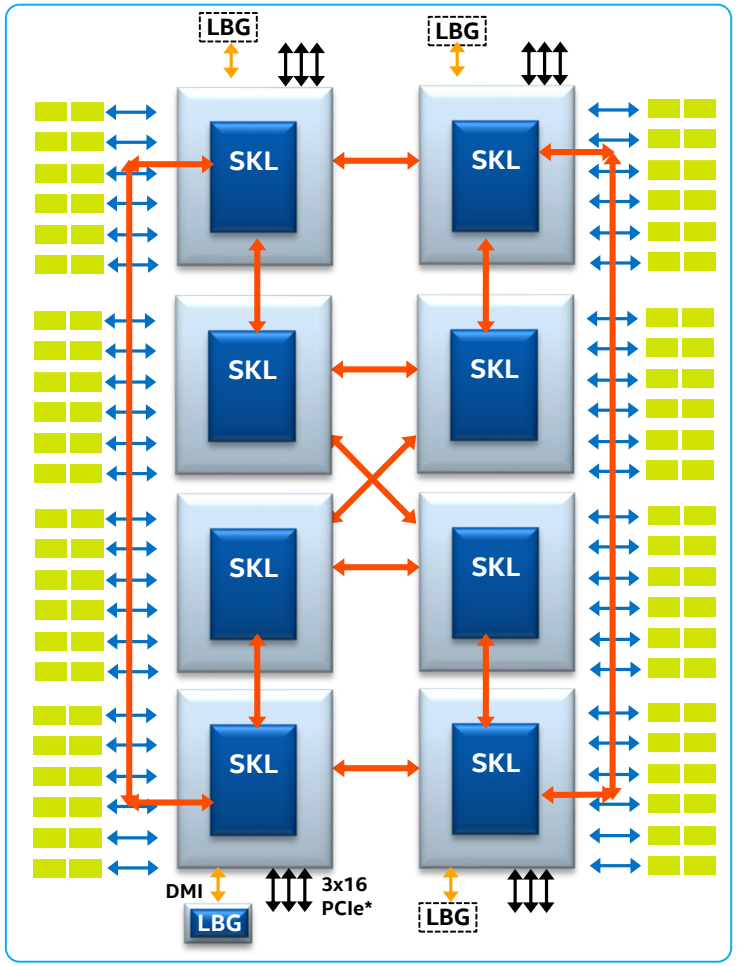
(2S-2UPI & 2S-3UPI shown)

4S Configurations



(4S-2UPI & 4S-3UPI shown)

8S Configuration



INTEL® XEON® SCALABLE PROCESSOR SUPPORTS CONFIGURATIONS RANGING FROM 2S-2UPI TO 8S

Intel® Advanced Vector Extensions 512 (Intel® AVX-512)

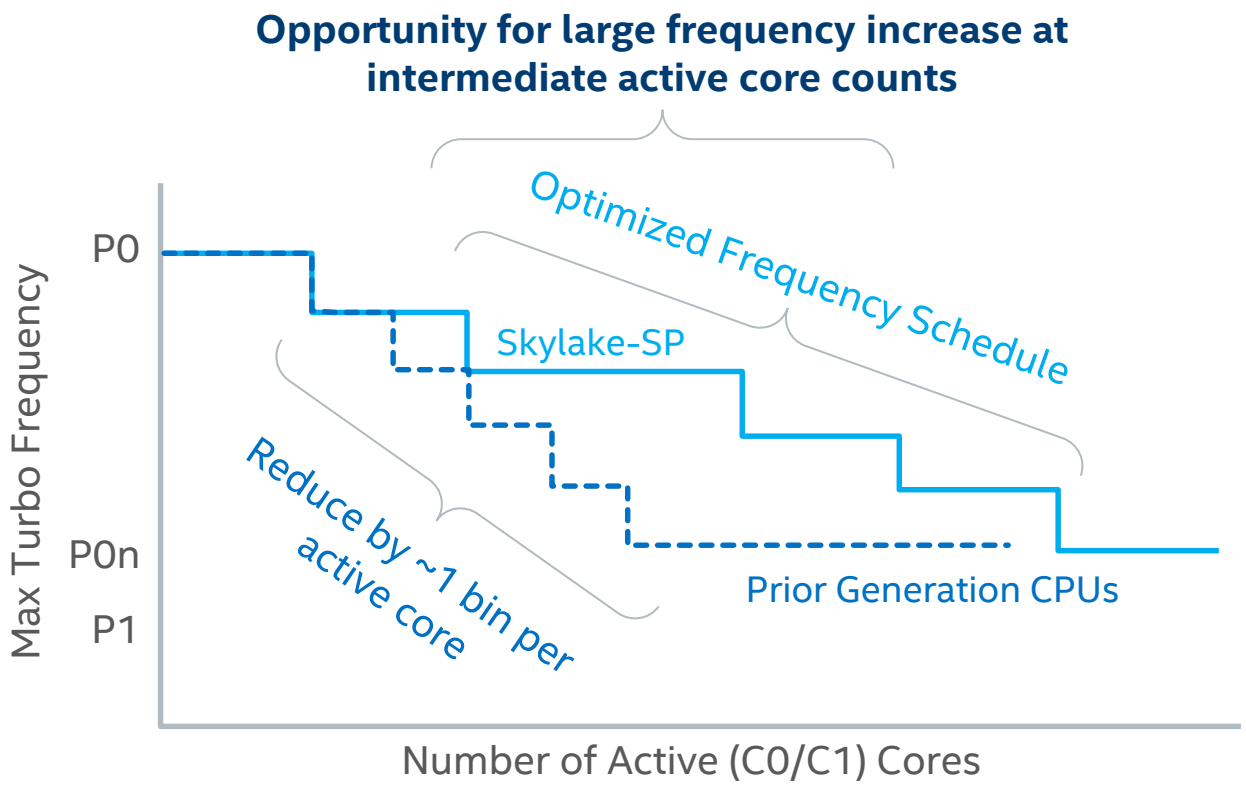
- 512-bit wide vectors
- 32 operand registers
- 8 64b mask registers
- Embedded broadcast
- Embedded rounding

Microarchitecture	Instruction Set	SP FLOPs / cycle	DP FLOPs / cycle
Skylake	Intel® AVX-512 & FMA	64	32
Haswell / Broadwell	Intel AVX2 & FMA	32	16
Sandybridge	Intel AVX (256b)	16	8
Nehalem	SSE (128b)	8	4

Intel AVX-512 Instruction Types	
AVX-512-F	AVX-512 Foundation Instructions
AVX-512-VL	Vector Length Orthogonality : ability to operate on sub-512 vector sizes
AVX-512-BW	512-bit Byte/Word support
AVX-512-DQ	Additional D/Q/SP/DP instructions (converts, transcendental support, etc.)
AVX-512-CD	Conflict Detect : used in vectorizing loops with potential address conflicts

POWERFUL INSTRUCTION SET FOR DATA-PARALLEL COMPUTATION

Optimized Turbo Profiles



Prior generation data center CPUs typically decreased turbo by 1 bin for each additional active core

Skylake-SP provides higher intermediate turbo points by stepping down in a more optimal manner

- Higher performance dynamically with C-states
- BIOS/OS core disable can be used to mimic higher frequency SKUs (with some tradeoffs)

Note: there is no guarantee that these frequencies can be achieved for a given workload on all units

*Picture is an illustration only. Not intended to represent any specific SKU or imply any frequency commitments.

Skylake-SP with Integrated Fabric

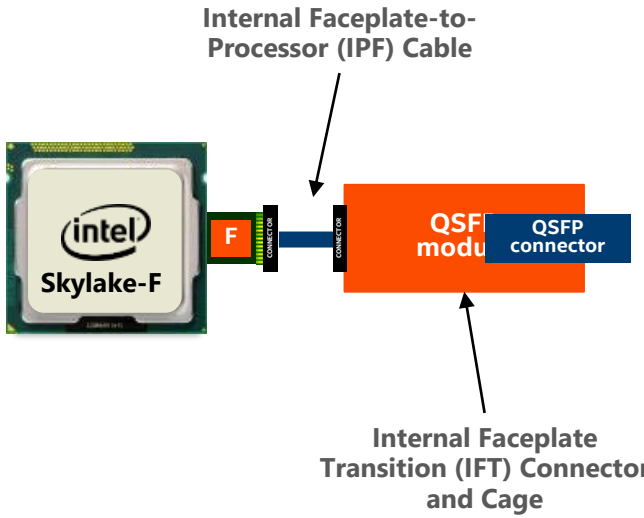
Single on-package Omni-Path Host Fabric Interface (HFI)

Fabric component interfaces to CPU using x16 PCIe* lanes

Fabric PCIe lanes are additional to the 48 PCIe lanes on the socket

Single cable from SKL-F package connector to QSFP module

Same socket for Skylake-SP and Skylake-F processors



- Purley platform can be designed to support both processors
- Platform design requires an expanded keep-out zone and additional board components to accommodate both processors

Intel® Xeon Phi™ Processor – TCO Solution for HPC & AI

A Key Element of HPC, AI, and Mixed Workload Clusters

Total Cost of Ownership

- Price Performance
- Power Efficiency
- Performance

Optimized for HPC & AI

- Highly-Parallel
- No PCIe Bottlenecks
- Scalability

Complements Intel® Xeon®

- Common Programming
- Mixed Clusters
- Runs x86 code

Reduces total cost of ownership, designed for HPC & AI, protects investment

Intel® Xeon Phi™ Processor Architecture

Self-Boot Processor
Binary-compatibility with Xeon, 3+ TFLOPS[1] (DP)

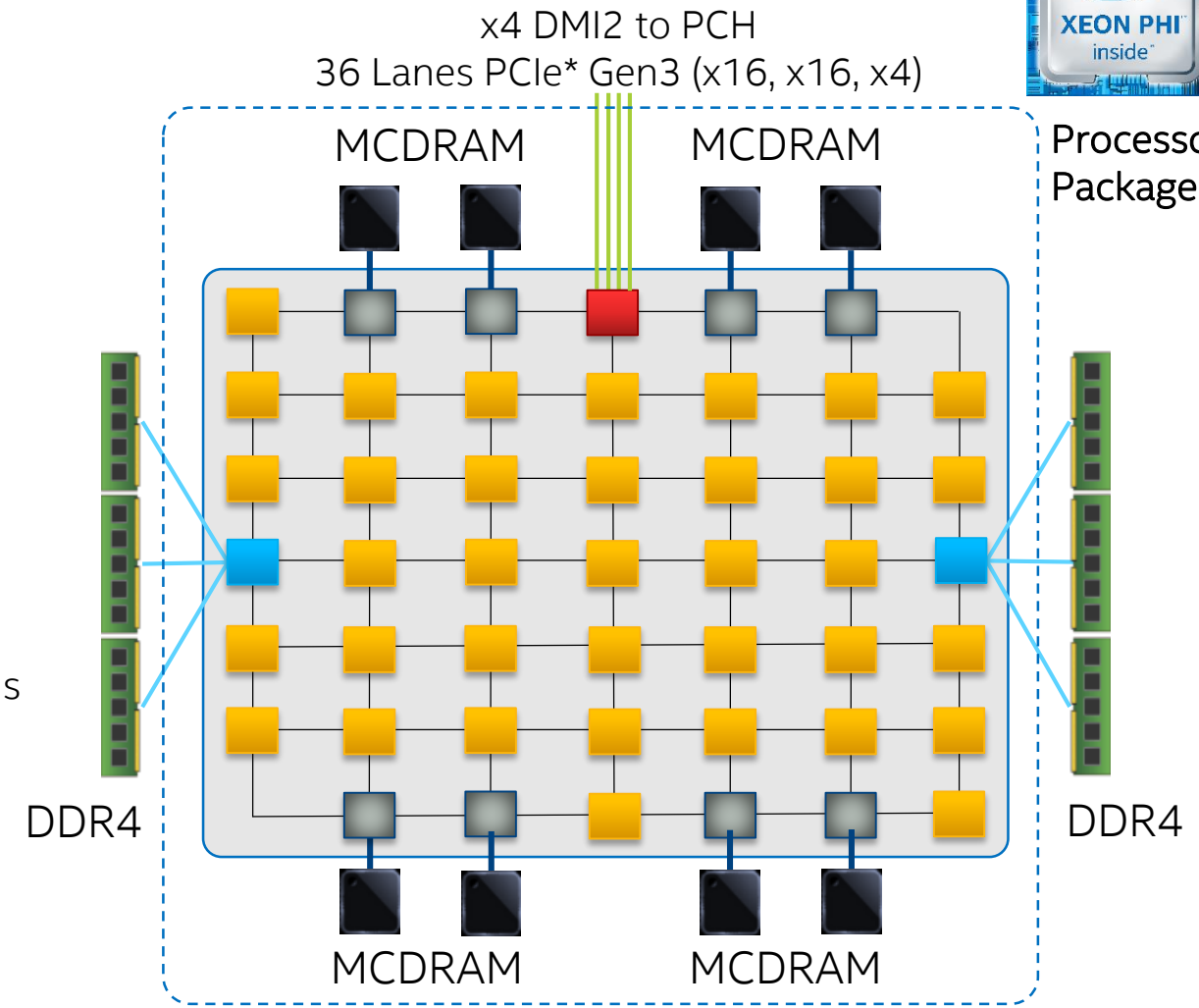
On-package memory
16GB, up to 490 GB/s STREAM TRIAD

Platform Memory
Up to 384GB (6ch DDR4-2400 MHz)

Other Key Features
- ✓ 2D Mesh Architecture
- ✓ Out-of-Order Cores
- ✓ 3X Single-Thread vs. KNC
- ✓ Intel® AVX-512 Instructions
- ✓ Scatter/Gather Engine
- ✓ Integrated Fabric - OPA

TILE: (up to 36)

Enhanced Intel® Atom™ cores based on Silvermont Microarchitecture

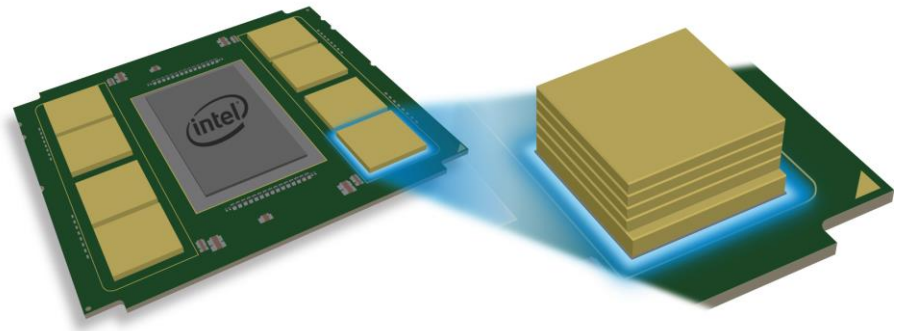


Tile | EDC (Embedded DRAM Controller) | IMC (Integrated Memory Controller) | IIO (Integrated I/O Controller)

[1]Theoretical peak performance

Bringing Memory Back Into Balance

up to 16 GB of High Bandwidth on-package memory in Knights Landing



3 Modes of Operation:

Flat Mode: Acts as Memory

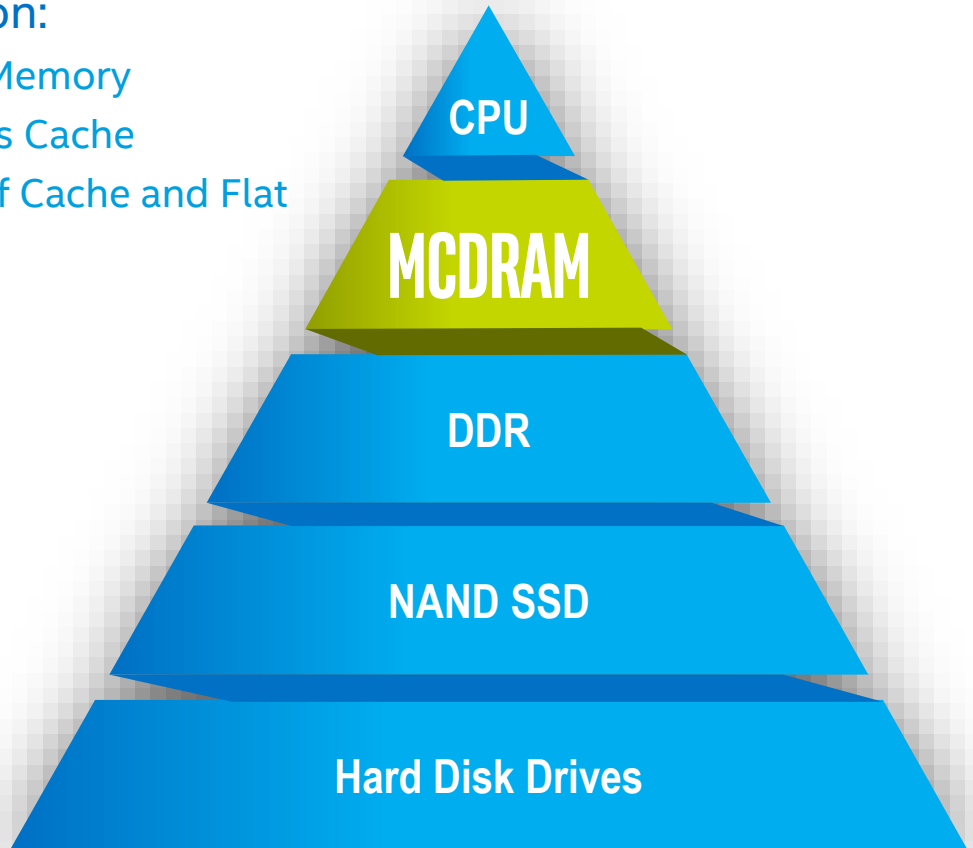
Cache Mode: Acts as Cache

Hybrid Mode: Mix of Cache and Flat

5X
BANDWIDTH
VS. DDR4[1],
>400 GB/s[1]

>5X
ENERGY EFFICIENT
VS. GDDR5

>3X
DENSITY
VS. GDDR5[2]



[1] Projected result based on internal Intel analysis of STREAM benchmark using a Knights Landing processor with 16GB of ultra high-bandwidth versus DDR4 memory with all channels populated.
[2] Projected result based on internal Intel analysis comparison of 16GB of ultra high-bandwidth memory to 16GB of GDDR5 memory used in the Intel® Xeon Phi™ coprocessor 7120P.

Intel® Xeon Phi™ Product Family x200

Intel® Xeon Phi™ Processor

with integrated
Intel® Omni-Path Fabric

Host Processor in Groveport Platform

Self-boot Intel® Xeon Phi™ processor

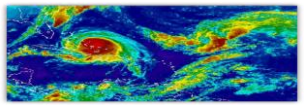
Intel® Xeon Phi™ Target Segments & Applications

Deep Learning Training

Material Science: **VASP*, NWCHEM*, GTC-P***

QCD: **QPHIX*, MILC*, CHROMA*, CCS QCD***

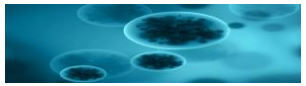
CFD/Mfg: **OPENFOAM*, CLOVERLEAF*, LSTC LS-DYNA*, CONVERGENT SCIENCE CONVERGE CFD***



Weather/Climate/Cosmology: **WRF*, NEMO*, WALLS***

Energy: **ISO3DFD***

FSI: **STAC A2*, MONTE CARLO*, BLACK SCHOLES*, BINOMIAL OPTIONS***



MD: **LAMMPS*, NAMD*, GROMACS*, AMBER***

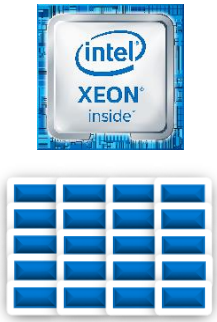
Features Driving Perf & Perf/$/W
16GB MCDRAM
High memory (MCDRAM) BW (≤ 490 GB/s)
Intel® AVX-512 ER
High system memory (≤ 400 GB)
High number of physical cores (≤ 72)
High number of threads (≤ 288)
Lower system price (~$4700)[1]
Lower system price (~$4700)[1]

*Other names and brands may be claimed as the property of others.

[1]Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to www.intel.com/benchmarks. Configurations: See Slides 40-52.

Intel® Xeon Phi™ Utilization Value

Homogenous "Large" Core

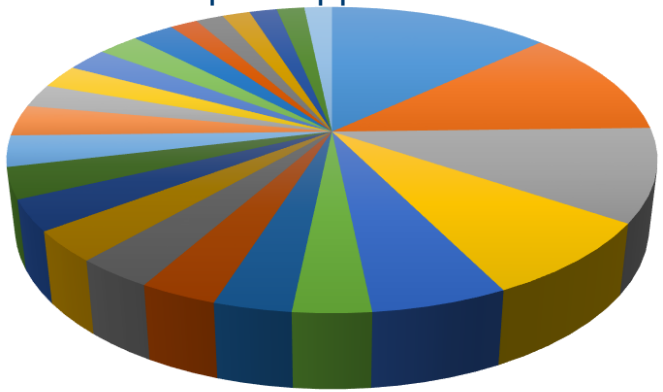


Homogenous "Small" Core

General Purpose ⟷ **Workload Optimized**

Intel® Xeon Phi™ Utilization Benefits
Runs optimized applications best
Runs all x86 applications
Doesn't reduce resources for some applications

Example Supercomputer Cluster
Top 25 Applications



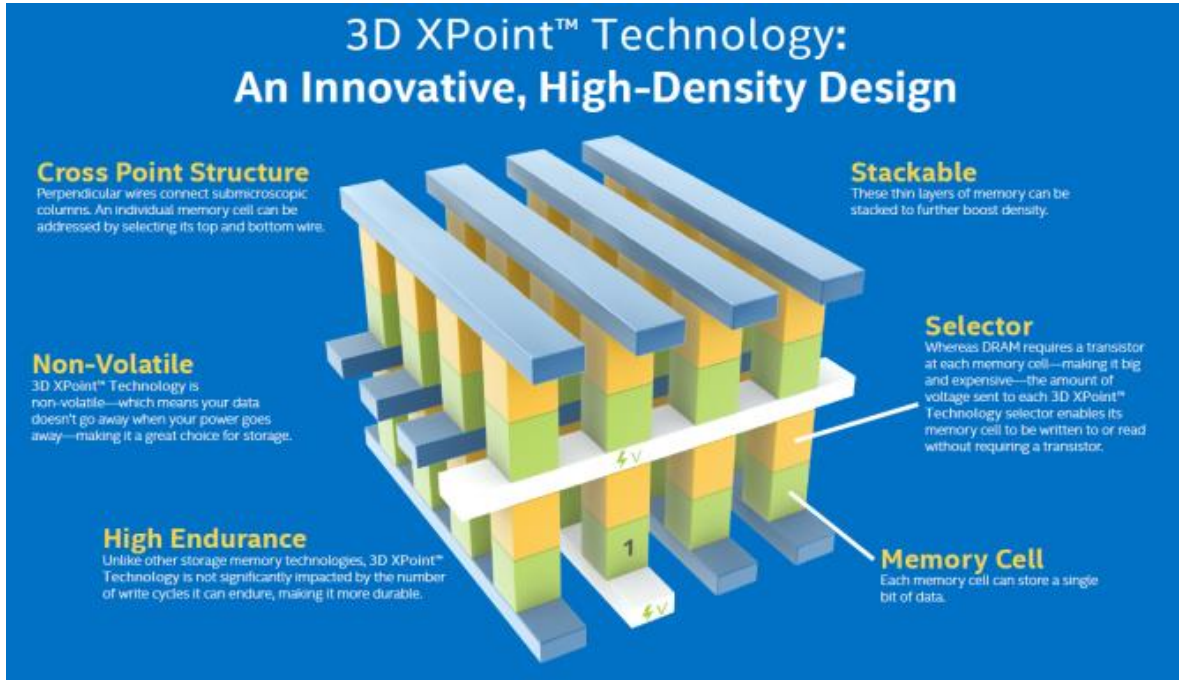
GPU Utilization Limitations
Requires coding to run application, requires optimization to run best
Doesn't run x86 applications
Dedicated resource reduces cluster performance for some applications

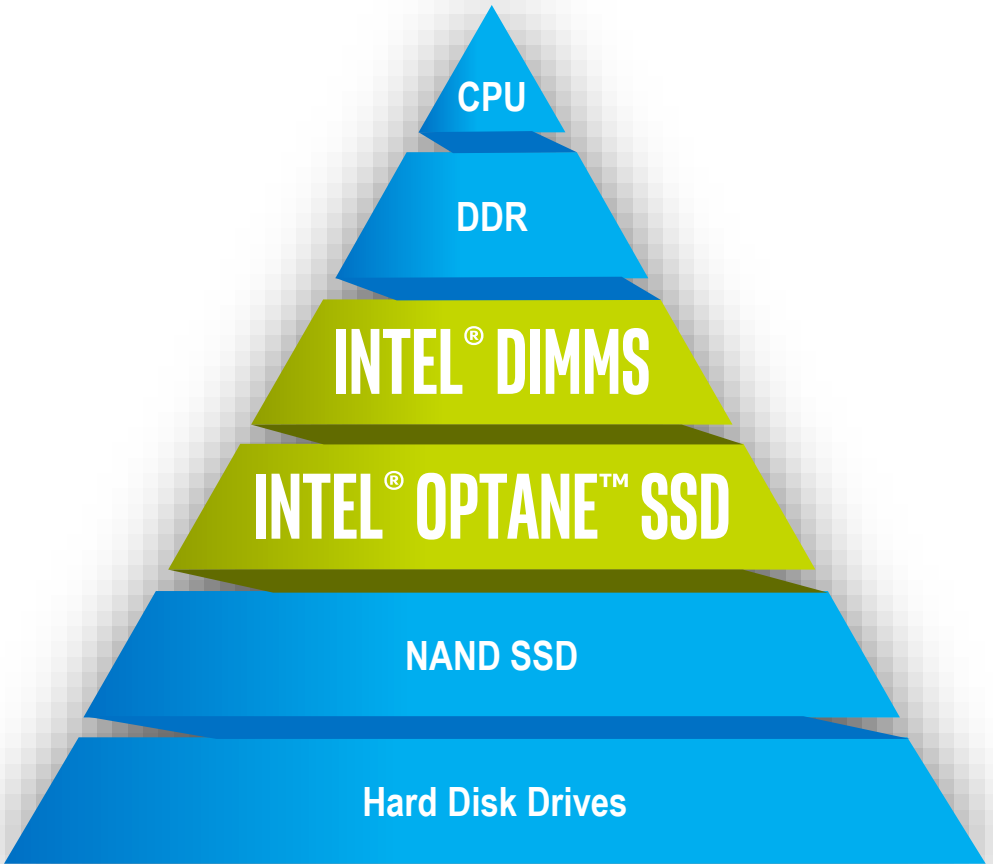
Bridging the Memory-Storage Gap

Memory

Intel® Scalable System Framework

Intel® Optane™ Technology Based on 3D XPoint™





SSD

- 10x More Dense than Conventional Memory[3]
- Intel® Optane™ SSDs 5-7x Current Flagship NAND-Based SSDs (IOPS)[1]

DRAM-like performance

- Intel® DIMMs Based on 3D-XPoint™
- 1,000x Faster than NAND[1]
- 1,000x the Endurance of NAND[2]

[1] Performance difference based on comparison between 3D XPoint™ Technology and other industry NAND
[2] Density difference based on comparison between 3D XPoint™ Technology and other industry DRAM
[2] Endurance difference based on comparison between 3D XPoint™ Technology and other industry NAND

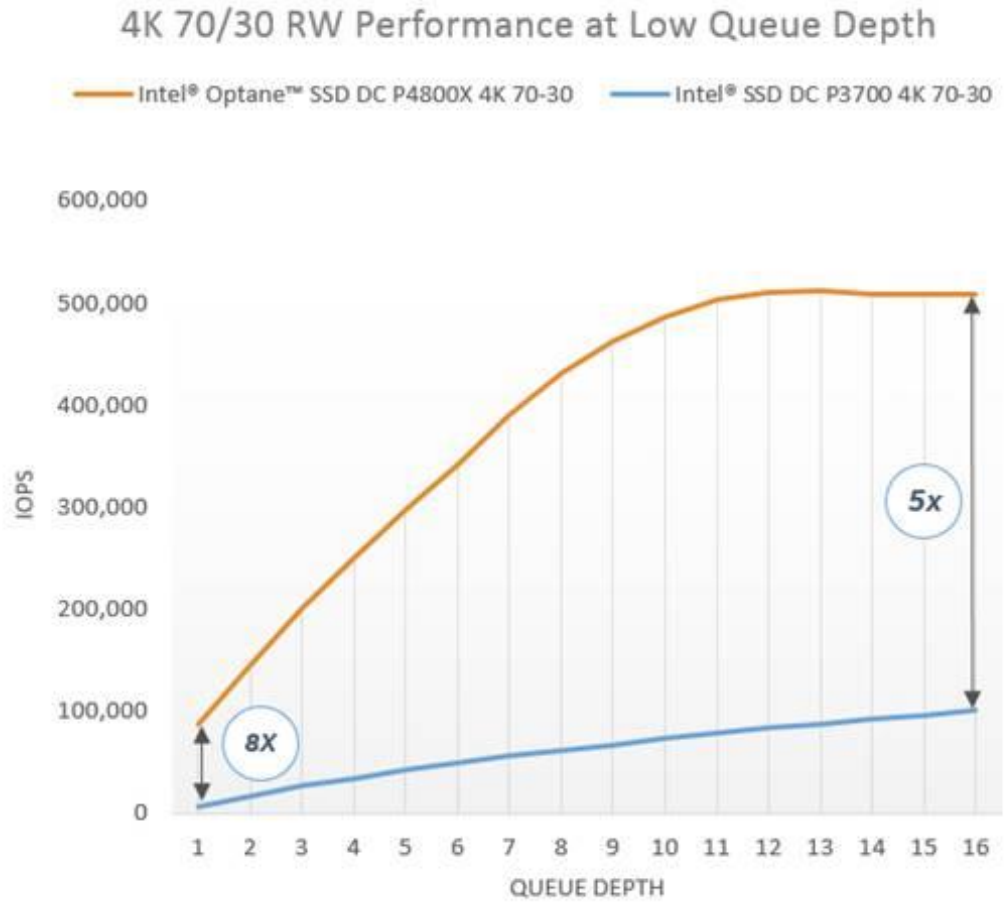
intel

World's Most Responsive Data Center SSD[1]

Delivering an **industry leading combination of low latency, high endurance, QoS and high throughput**, the Intel® Optane™ SSD is the first solution to **combine the attributes of memory and storage**. This innovative solution is optimized to **break through storage bottlenecks** by providing a new data tier. It accelerates applications for **fast caching and storage, increasing scale per server** and reducing transaction cost. Data centers based on the latest Intel® Xeon® processors can now also **deploy bigger and more affordable datasets** to gain new insights from larger memory pools.

1. Responsiveness defined as average read latency measured at Queue Depth 1 during 4k random write workload. Measured using FIO 2.15. Common configuration - Intel 2U Server System, OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Intel drives evaluated - Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. Samsung* drives evaluated – Samsung SSD PM1725a, Samsung SSD PM1725, Samsung PM963, Samsung PM953. Micron* drive evaluated – Micron 9100 PCIe* NVMe* SSD. Toshiba* drives evaluated – Toshiba ZD6300. Test – QD1 Random Read 4K latency, QD1 Random RW 4K 70% Read latency, QD1 Random Write 4K latency using FIO 2.15.

Breakthrough Performance



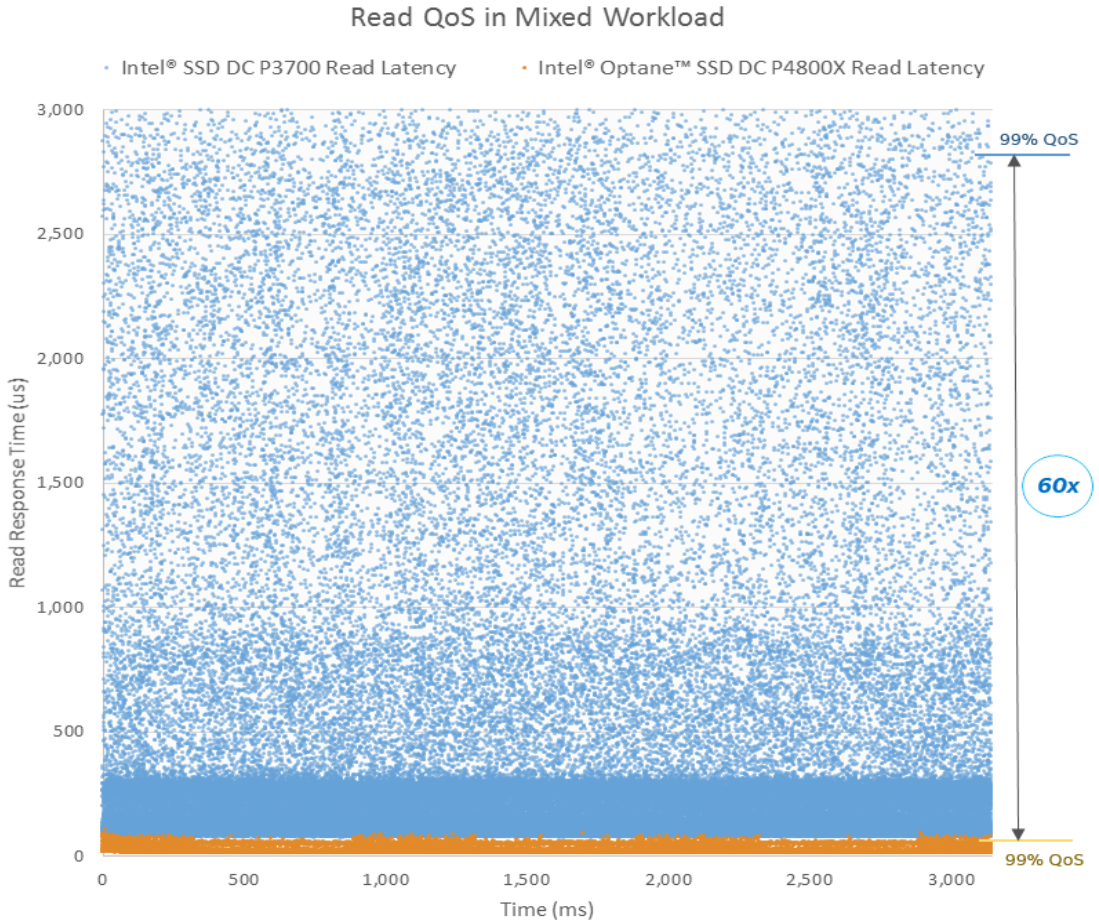
intel XEON inside + intel OPTANE™

- **5-8x faster** at low Queue Depths[1]
- Vast majority of **applications generate low QD** storage workloads

1. Common Configuration - Intel 2U PCSD Server ("Wildcat Pass"), OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. Performance – measured under 4K 70-30 workload at QD1-16 using fio-2.15.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance.

Predictably Fast Service



intel XEON inside + intel OPTANE™

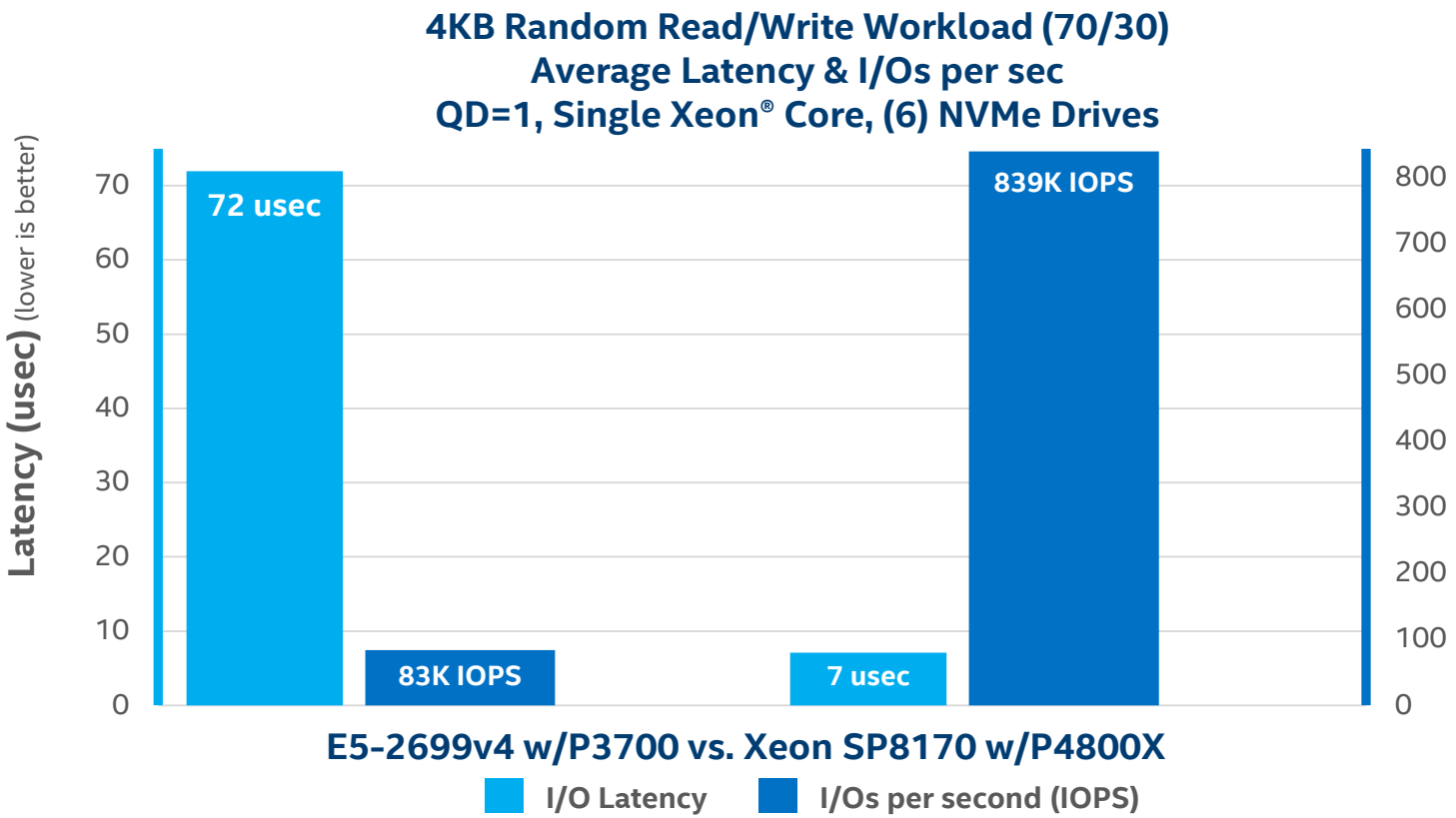
- ☑ up to **60X** better at 99% QoS[1]
- ☑ Ideal for critical applications with aggressive latency requirements

1. Common Configuration - Intel 2U PCSD Server ("Wildcat Pass"), OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. QoS – measures 99% QoS under 4K 70-30 workload at QD1 using fio-2.15.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance.

Intel® Optane™ SSD DC P4800X for Storage Builders

SPDK Performance: Platform Comparison

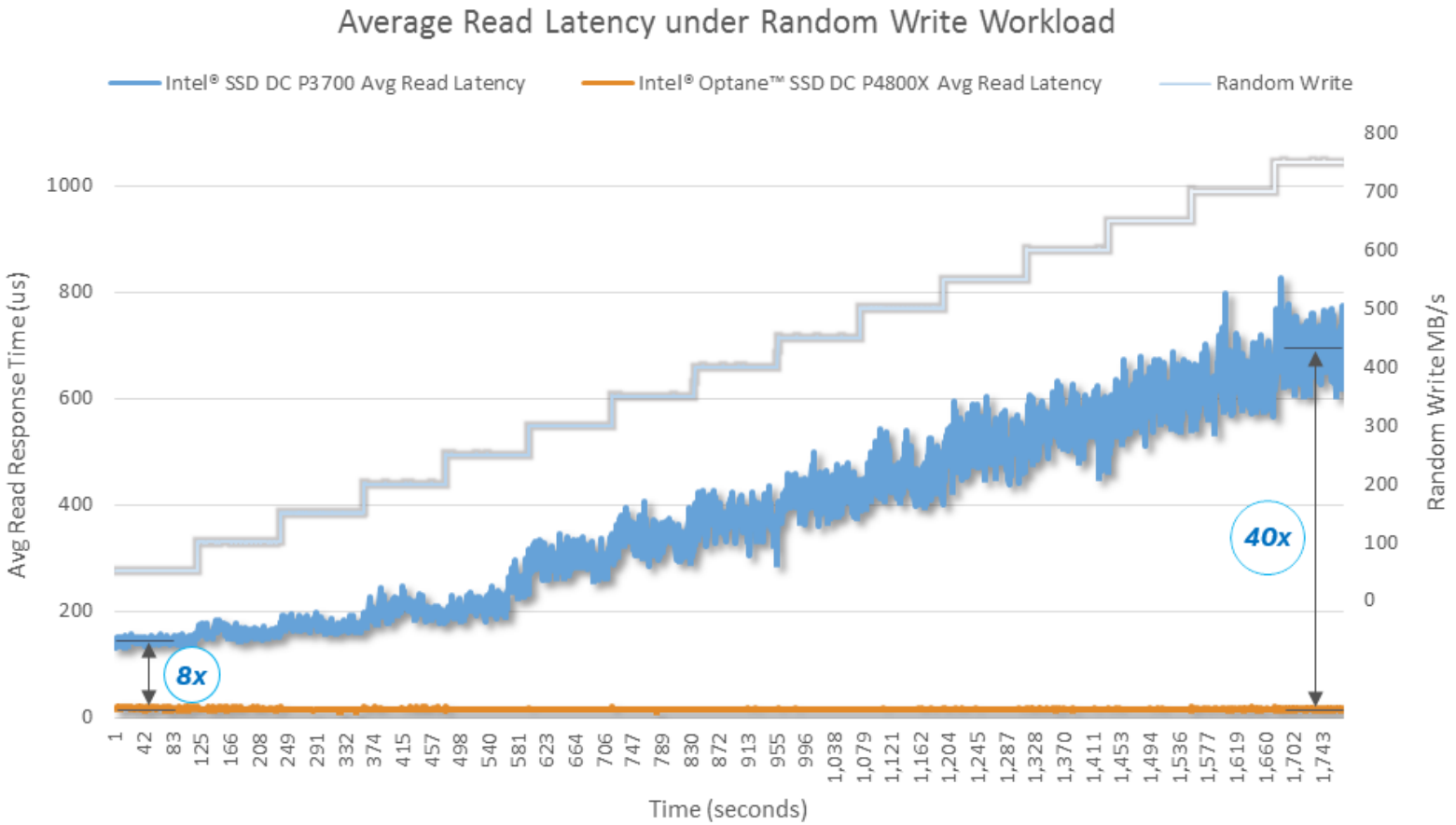


Intel® Xeon® Scalable Processor Platinum Family + Intel® Optane™

- 10X higher throughput
- 10X lower latency
- Up to 27 cores remaining for:
  - Virtual Machines
  - Big Data/Analytics
  - Machine Learning
  - Storage services like erasure coding, de-duplication, compression, or encryption.
- Platform offers RDMA
  - Enables NVMe over Fabrics
  - No more trapped I/O capacity

See notices, configurations, disclaimers

Responsive Under Load



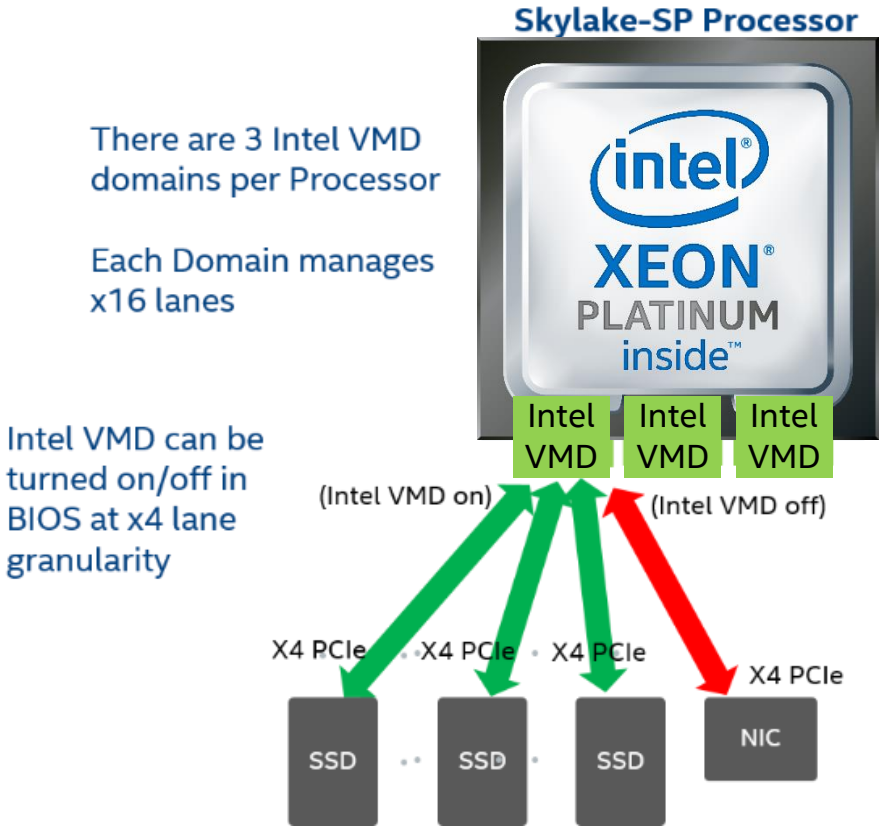
+ intel OPTANE™

- ☑ up to **40X faster response time** under workload[1]
- ☑ Consistently **amazing response time under load**

1. Responsiveness defined as average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 2.15. Common Configuration - Intel 2U PCSD Server ("Wildcat Pass"), OS CentOS 7.2, kernel 3.10.0-327.el7.x86_64, CPU 2 x Intel® Xeon® E5-2699 v4 @ 2.20GHz (22 cores), RAM 396GB DDR @ 2133MHz. Configuration – Intel® Optane™ SSD DC P4800X 375GB and Intel® SSD DC P3700 1600GB. Latency – Average read latency measured at QD1 during 4K Random Write operations using fio-2.15.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance.

Intel® Volume Management Device (Intel® VMD)

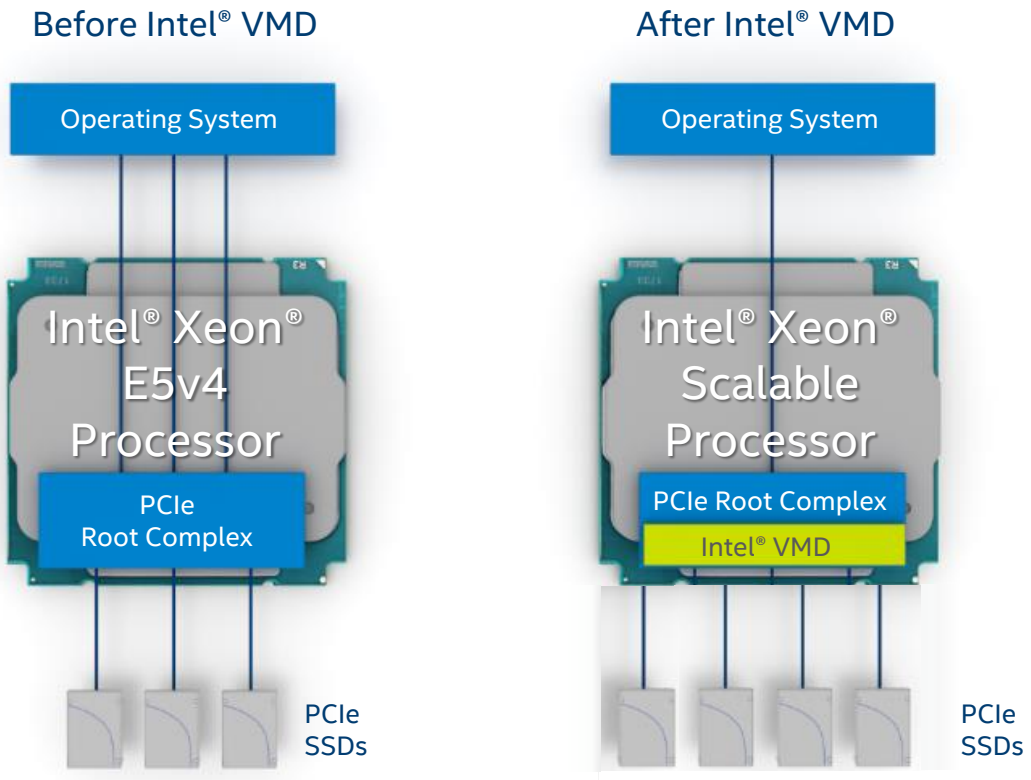


Intel® VMD is a CPU-integrated device to aggregate NVMe SSDs into a storage volume and enables other storage services such as RAID

- Intel® VMD is an "integrated end point" that stops OS enumeration of devices under it
- Intel® VMD maps entire PCIe* trees into its own address space (a domain)
- Intel® VMD driver sets up and manages the domain (enumerate, event/error handling), but out of fast IO path

ELIMINATES ADDITIONAL COMPONENTS TO PROVIDE A FULL-FEATURE STORAGE SOLUTION

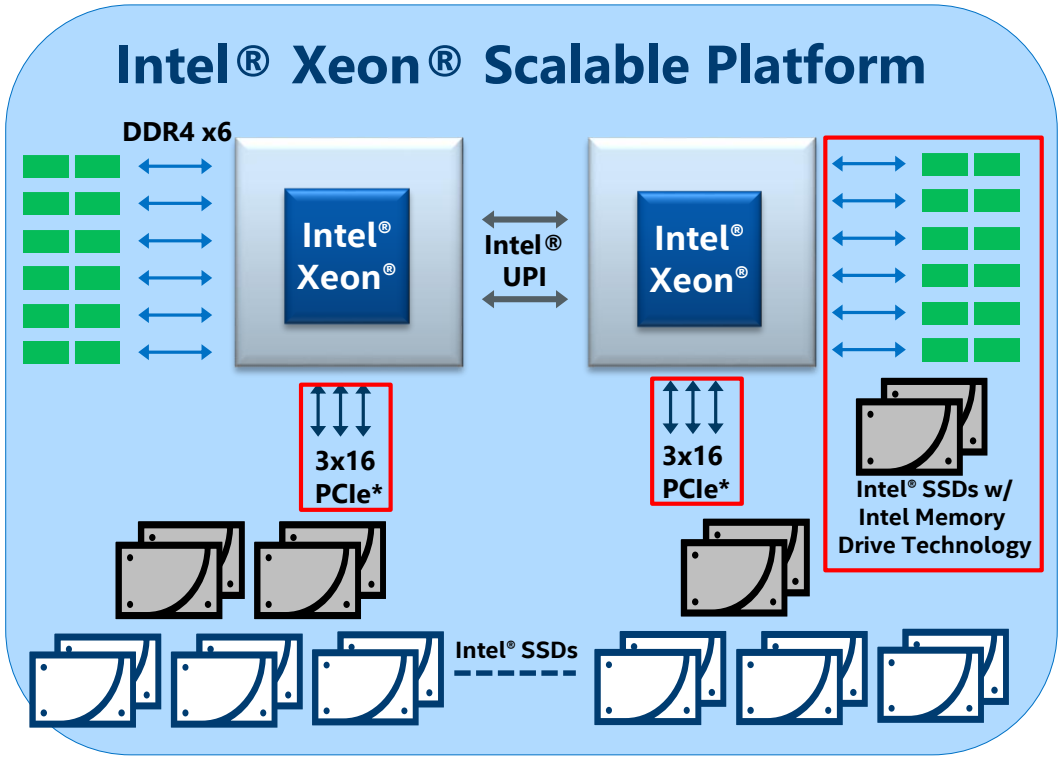
Intel® Volume Management Device



- Intel® VMD is a new technology to enhance solutions with PCIe* storage
- Supported for Windows, Linux, and ESXi*
- Multi-SSD vendor support
- Intel® VMD enables:
  - Isolating fault domains for device surprise hot-plug and error handling
  - Provide consistent framework for managing LEDs
  - Simplify PCIe storage software stacks

Intel® VMD enables customers to **simplify and harden** solutions using PCIe storage.

Scale up or out with more PCIe lanes, Intel® SSDs, and Intel Memory Drive Technology

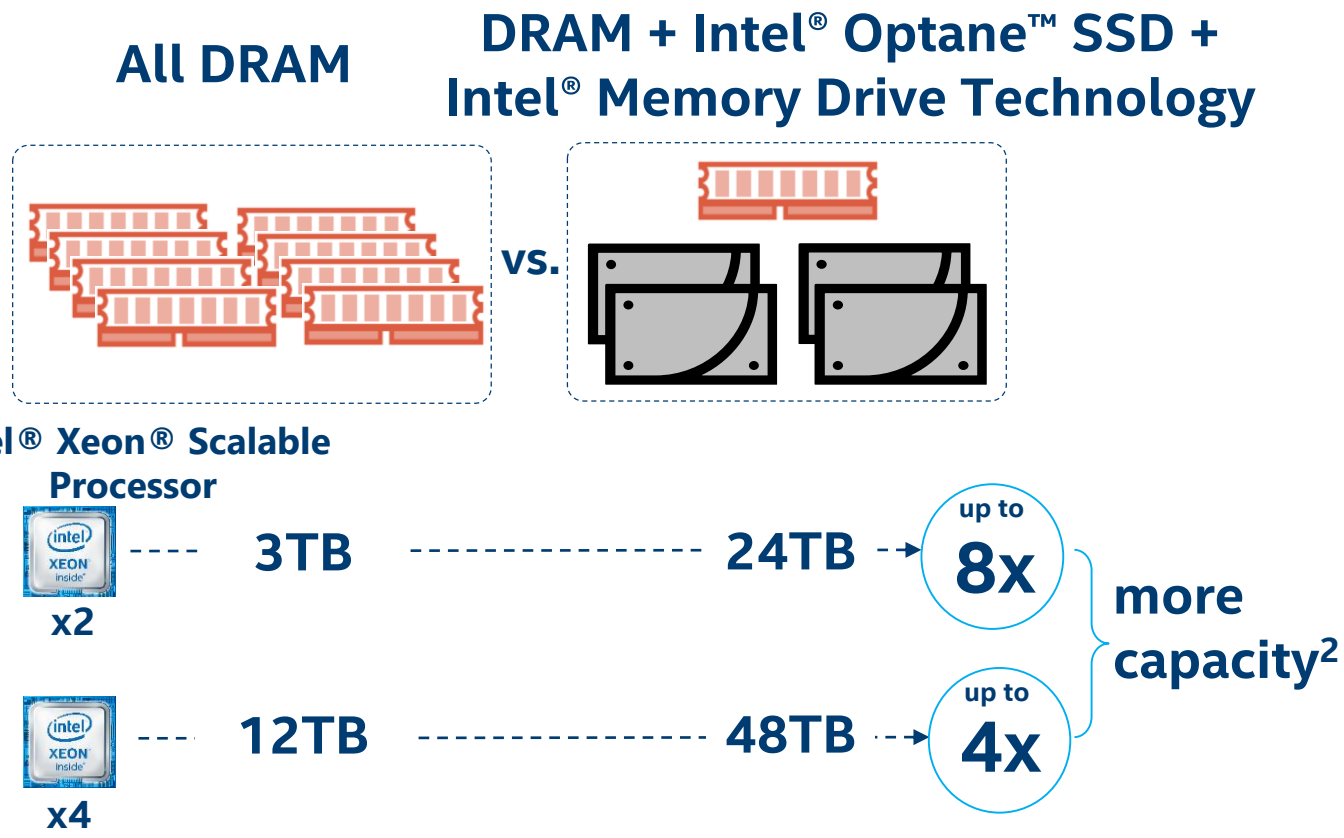


Scale up memory with Intel® Optane™ SSD and Intel Memory Drive Technology

- integrates **transparently into memory** subsystem with no OS or app changes[1]
- DRAM + Intel® Optane™ SSD + Intel® Memory Drive Technology emulate a **single volatile memory pool**

Scale out capacity and performance with **20% more PCIe lanes[2]** and a broad portfolio of Intel® SSDs

Massively Scalable, Faster[3] Memory Pools



- Increase memory pool up to 8x[1]
- Displace DRAM up to 10:1 in select workloads[2]
- Higher platform memory & PCIe bandwidth with Intel® Scalable Processor[3]
- Accelerate applications and gain new insights from larger working sets

See notices, configurations, disclaimers

Thanks!

nikolay.mester@intel.com