

INTEL® XEON® SCALABLE PLATFORM

Notices and Disclaimers

Intel technologies may require enabled hardware, specific software, or services activation. Performance varies depending on system configuration. Check with your system manufacturer or retailer.

For more complete information about performance and benchmark results, visit <u>www.intel.com/benchmarks</u>.

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.

Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction. § For more information go to http://www.intel.com/performance.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

No computer system can be absolutely secure.

Intel, the Intel logo, Xeon, Intel vPro, Intel Xeon Phi, Look Inside., are trademarks of Intel Corporation in the U.S. and/or other countries.

© 2017 Intel Corporation.

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



CLOUD

CLOUD ECONOMICS

AI & ANALYTICS

INTELLIGENT DATA PRACTICES

5G

NETWORK TRANSFORMATION

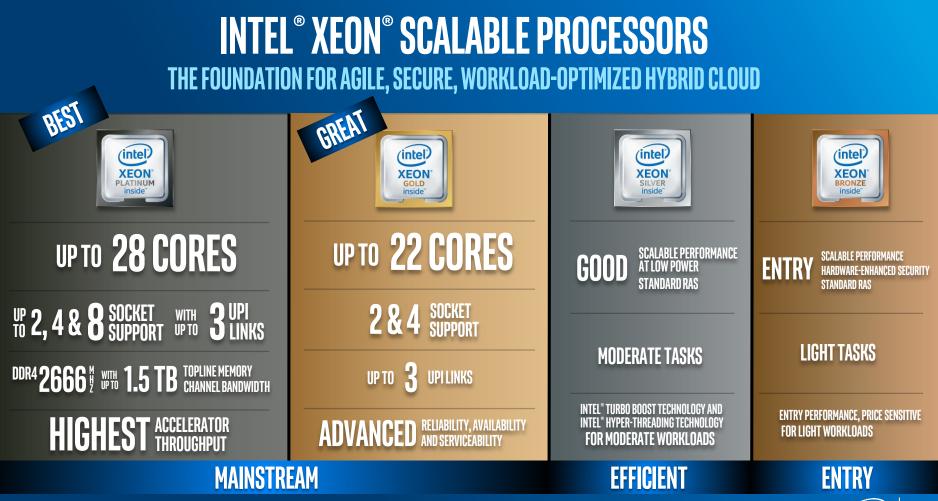
MEGATINENDS

TRANSFORMATION STARTS ON THE INSIDE

INTEL® XEON® SCALABLE PLATFORM



THE INDUSTRY'S BIGGEST PLATFORM ADVANCEMENT IN A DECADE



intel

THE INTEL® XEON® SCALABLE PROCESSORS: AN AGILE SOLUTION STACK FOR DATA CENTER WORKLOADS

41XX Ser

INTEL® XEON® BRONZE PROCESSOR

31XX Series Bronze (2 Socket)

- Up to 8 cores
- 2S configuration
- Improved core interconnect (UPI)
 over past gen
- 48 PCIe 3.0 lanes
- Intel® AVX-512 feature enabled
- Standard RAS features

Entry Performance and security for price sensitive deployments

INTEL° XEON° SILVER PROCESSOR

41XX Series Silver (2 Socket)

- + Up to 12 cores
- + 2S configuration with Improved Memory channel performance
- + Intel[®] Turbo Boost Technology for higher frequency capability
- + Intel[®] HT Technology for hyper threaded workloads

Efficient Performance at Low

for single purpose workloads

Power. Provides more horsepower

INTEL[®] XEON[®] GOLD PROCESSOR

61XX Series Gold (2 and 4 Socket)

- + Up to 22 Cores
- + Added 3rd UPI link for increased dataflow across cores
- Increased performance across memory channels⁶
- + Intel[®] AVX-512 with additional FMA
- + Added Node Controller Support to assist in scaled node management

51XX Series Gold (2 and 4 Socket)

- + Up to 14 cores
- + Supports **2S and 4S configuration** for increased scalability
- + Increased core interconnect speed to boost data flow in multi-processor workloads
- + Advanced RAS features

Mainstream Performance, Fast Memory, More Interconnect Engines, Advanced Reliability

INTEL[®] XEON[®] PLATINUM PROCESSOR

81XX Series Platinum (2, 4, and 8 Socket)

- + Up to 28 Cores
- + 2,4, or 8 socket configurations for best performance and scalability⁵
- + Topline memory channel performance (1.5 TB memory bandwidth on select SKUs)
- + 3 UPI links option across 2S,4S,8S for improved scalability and intercore data flow

The Best Performance, Scalability, Core options, and all Hardware-Enhanced Security features for the most robust capability

Better performance, interconnectivity, scalability, and memory

Unified Intel® Xeon® Scalable Platform



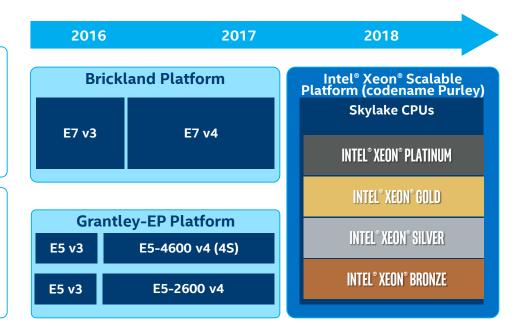
Intel[®] Xeon[®] Processor E7

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/\$



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	6 Channels DD	R4	
Cores Per Socket	Up to 22	Up to 28	DDR4	Core	Core
Threads Per Socket	Up to 44 threads	Up to 56 threads	DDR4		
Last-level Cache (LLC)	Up to 55 MB	Up to 38.5 MB (non-inclusive)	DDR4	Core	Core
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 / PCle* 3.0 (2.5, 5, 8 GT/s)	48 / 12 / PCIe 3.0 (2.5, 5, 8 GT/s)	DDR4	Core	Core ed L3
Memory Population	4 channels of up to 3 RDIMMs, LRDIMMs, or 3DS LRDIMMs	6 channels of up to 2 RDIMMs, LRDIMMs, or 3DS LRDIMMs	DDR4	Omni-P	ath HF
Max Memory Speed	Up to 2400	Up to 2666	48 Lanes PCIe* 3.0		
TDP (W)	55W-145W	70W-205W			

2 or 3 UPI UPI

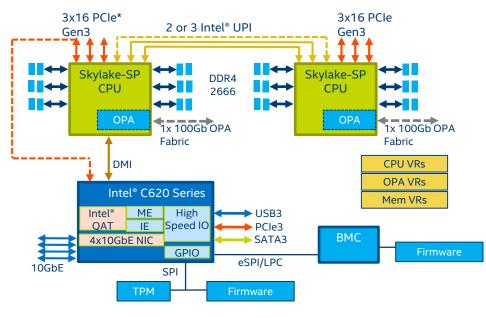
UPI

UPI

Omni-Path

DMI3

Platform Overview



BMC: Baseboard Management Controller		PCH: Intel® Platform Controller Hub	IE: Innovation Engine	
Intel® OPA: Intel® Omni-Path Architecture		Intel QAT: Intel® QuickAssist Technology	ME: Manageability Engine	
	NIC: Network Interface Controller	VMD: Volume Management Device	NTB: Non-Transparent Bridge	

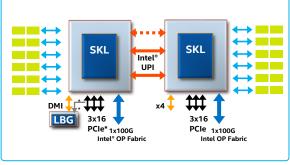
Capabilities	Details
Socket	Socket P
Scalability	2S, 4S, 8S, and >8S (with node controller support)
CPU TDP	70W – 205W
Chipset	Intel® C620 Series (code name Lewisburg PCH)
Networking	Intel® Omni-Path Fabric (integrated w/ CPU + discrete) 4x10GbE (integrated w/ chipset) 100G/40G/25G discrete options
Compression and Crypto Acceleration	Intel® QuickAssist Technology option in chipset to support 100Gb/s comp/decomp/crypto 100K RSA2K public key
Storage	CPU integrated QuickData Technology, VMD, and NTB Intel® Optane™ SSD, Intel® 3D-NAND NVMe* & SATA SSD
Security	CPU Instruction Set enhancements (MBE, PPK, MPX) Manageability Engine with multiple secure boot options Intel® Platform Trust Technology Intel® Key Protection Technology
Manageability	Intel® Node Manager Intel® Datacenter Manager Innovation Engine (IE)



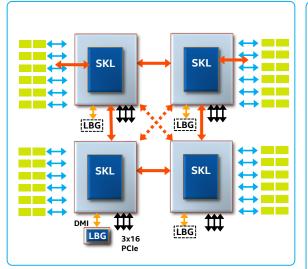
Platform Topologies

2S Configurations General Purpose/Storage/

HPC/Comms SP



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

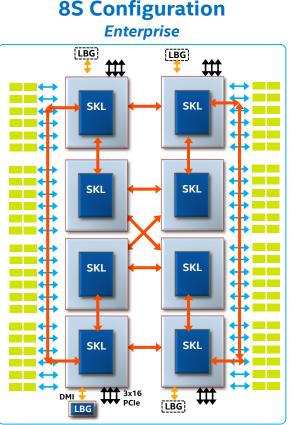


4S Configurations

Enterprise and Cloud

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

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





INTEL® XEON® SCALABLE PROCESSORS -Architectural enhancements

Overview

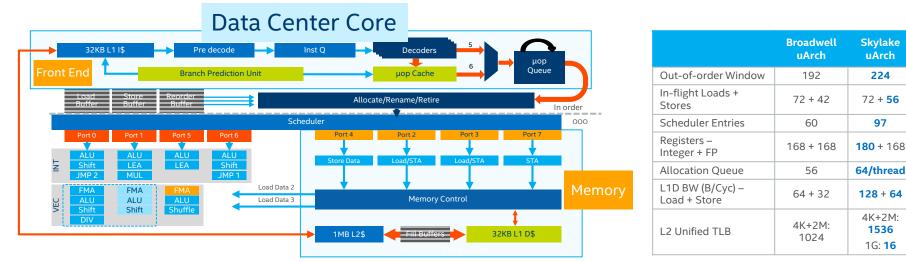
- Skylake core microarchitecture, with data center specific enhancements
- Intel[®] AVX-512 with 32 DP Flops per core
- Data center optimized cache hierarchy -1MB L2 per core, non-inclusive L3

- New mesh interconnect architecture
- Enhanced memory subsystem
- Modular IO subsystem with integrated devices
- New Intel[®] Ultra Path Interconnect (Intel[®] UPI

- Intel[®] Speed Shift Technology
- Security & Virtualization enhancements
- Optional Integrated Intel[®] Omni-Path Fabric (Intel[®] OPA)

Features	Intel® Xeon® Processor E5-2600 v4	Intel® Xeon® Scalable Processor (Skylake-SP)	6 Channels DI)R4		
Cores Per Socket	Up to 22	Up to 28	DDR4	Core	Core	2 or 3 UP
Threads Per Socket	Up to 44 threads	Up to 56 threads	DDR4	Core	Core	UPI
Last-level Cache (LLC)	Up to 55 MB	Up to 38.5 MB (non-inclusive)				UPI
QPI/UPI Speed (GT/s)	2x QPI channels @ 9.6 GT/s	Up to 3x UPI @ 10.4 GT/s	DDR4	Core	Core	
PCIe* Lanes/ Controllers/Speed(GT/s)	40 / 10 / PCle 3.0 (2.5, 5, 8 GT/s)	48 / 12 / PCIe 3.0 (2.5, 5, 8 GT/s)	DDR4	Share	ed L3	UPI Omni Path
Memory Population	4 channels of up to 3 RDIMMs, LRDIMMs, or 3DS LRDIMMs	6 channels of up to 2 RDIMMs, LRDIMMs, or 3DS LRDIMMs	DDR4 48 Lanes			DMI3
Max Memory Speed	Up to 2400	Up to 2666	PCIe 3.0		···· II	
TDP (W)	55W-145W	70W-205W		***1	• •	

Core Microarchitecture Enhancements



- Larger and improved branch predictor, higher throughput decoder, larger window
- Improved scheduler and execution engine, improved throughput and latency of divide/sqrt
- More load/store bandwidth, deeper load/store buffers, improved prefetcher
- Data center specific enhancements → Intel® AVX-512 with 2 FMAs per core, larger 1MB L2 per core

DATA CENTER-SPECIFIC ENHANCEMENTS TO THE CORE

Key Instruction Set Architecture Enhancements

COMPUTE

Intel[®] AVX-512

2x compute density per core for vector operations

Cache Management Instructions

CLFLUSHOPT – Lower latency cache line flush CLWB – Cache line writeback to memory without invalidation

VIRTUALIZATION

Improved Time Stamp Counter Virtualization

Reduces overhead on VMs migrating across processors running at different base frequency

SECURITY

Page Protection Keys (PPK)

Extends paging architecture to provide a page-granular, thread-private user-level memory protection

Mode Based Execution (MBE)

Protects against malicious kernel updates in a virtualized system



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

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

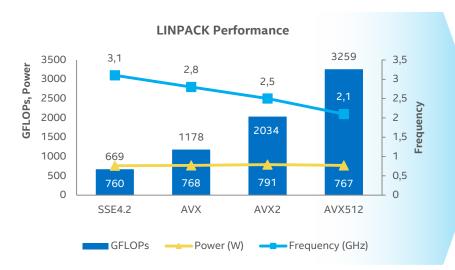
Microarchitecture	Instruction Set	SP FLOPs / cycle	DP FLOPs / cycle
Skylake	Intel® AVX-512 & FMA	64	32
Haswell / Broadwell	Intel [®] AVX2 & FMA	32	16
Sandybridge	Intel® AVX (256b)	16	8
Nehalem	SSE (128b)	8	4

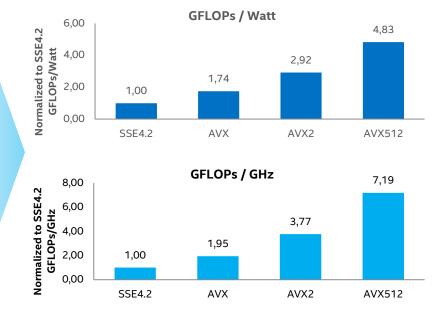
Intel AVX-512 Instruction Types			
Intel AVX-512-F	AVX-512 Foundation Instructions		
AVX-512-VL	Vector Length Orthogonality: ability to operate on sub-512 vector sizes		
AVX-512-BW	512-bit Byte/Word support		
AVX-512-DQ	Additional D/Q/SP/DP instructions (converts, transcendental support, etc.)		
AVX-512-CD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

INTEL® AVX-512 DOUBLES THE NUMBER OF FLOPS PER CYCLE



Performance and Efficiency with Intel® AVX-512





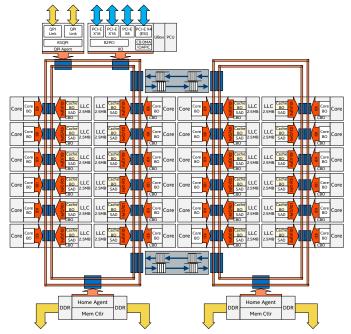
Intel® AVX is designed to balance power consumed by lowering frequency when needed, while delivering significant performance gains and reduced runtimes

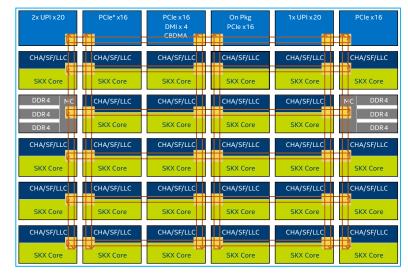
Source as of June 2017: Intel internal measurements on platform with Xeon Platinum 8180, Turbo enabled, UPI=10.4, SNC1, 6x32GB DDR4-2666 per CPU, 1 DPC. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark 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 products.

New Mesh Interconnect Architecture

Xeon® E7 v4 24-core die

Skylake-SP 28-core die



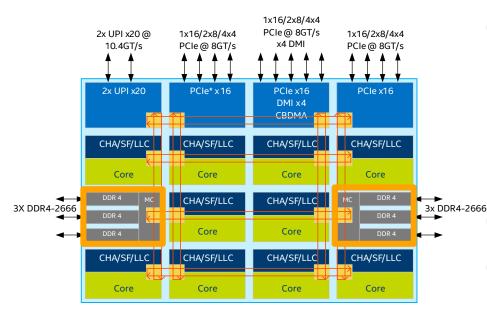


CHA – Caching and Home Agent ; SF – Snoop Filter; LLC – Last Level Cache; SKX Core – Skylake-SP Core; UPI – Intel® UltraPath Interconnect

DUAL-RING IN BROADWELL SERVER (INTEL® XEON® PROCESSOR E5/E7) V. MESH IN SKYLAKE-SP



Memory Subsystem



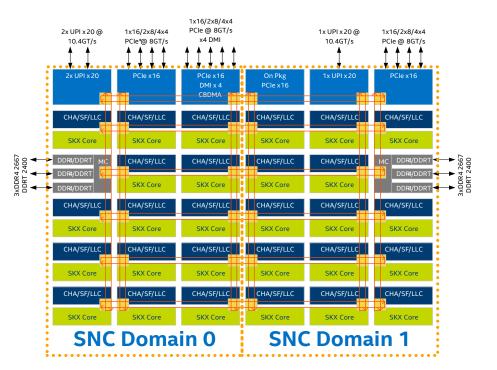
- 2 Memory Controllers, 3 channels each → total of 6 memory channels
 - DDR4 up to 2666, 2 DIMMs per channel
 - Support for RDIMM, LRDIMM, and 3DS-LRDIMM
 - 1.5TB Max Memory Capacity per Socket (2 DPC with 128GB DIMMs)
 - >60% increase in Memory BW per Socket compared to Intel[®] Xeon[®] processor E5 v4
- Supports various optimizations to reduce LLC miss latency
- Introduces a new memory device failure detection and recovery scheme - Adaptive Double Device Data Correction (ADDDC) - that reduces bandwidth and capacity overhead



Sub-NUMA Cluster (SNC)

- Prior generation supported Clusters-On-Die (COD)
- SNC provides similar localization benefits as COD, without some of its downsides:
 - One UPI caching agent even in 2-SNC mode
 - Latency for memory accesses in remote cluster is smaller, no UPI flow
 - LLC capacity is utilized more efficiently in 2cluster mode, no duplication of lines in LLC

	SNC	UPI Prefetch	
UMA	Disabled	Disabled	UMA, No prefetch
NUMA	Disabled	Enabled	Recommended default setting,1 cluster/socket
NUMA	Enabled	Enabled	2 clusters per socket



inte

Memory Performance

Bandwidth-Latency Profile

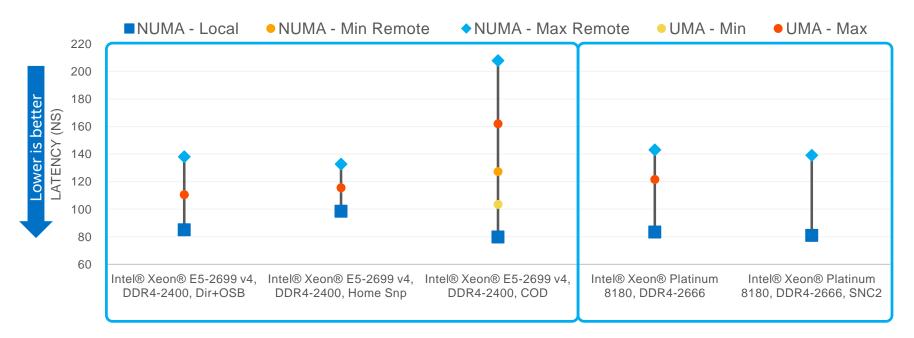


Source as of June 2017: Intel internal measurements on platform with Xeon Platinum 8180, Turbo enabled, UPI=10.4, SNC1/SNC2, 6x32GB DDR4-2400/2666 per CPU, 1 DPC, and platform with E5-2699 v4, Turbo enabled, 4x32GB DDR4-2400, RHEL 7.0. Software and workloads used in performance tests may have been optimized 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 information go to http://www.intel.com/performance



Memory Performance

Core to Memory Latency

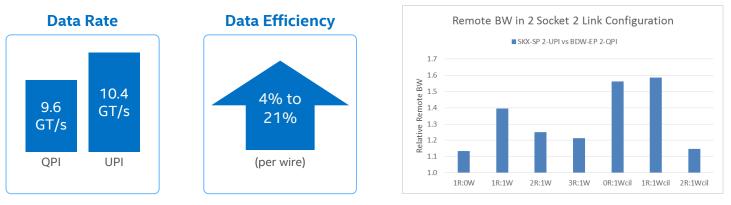


Source as of June 2017: Intel internal measurements on platform with Xeon Platinum 8180, Turbo enabled, UPI=10.4, 6x32GB DDR4-2666, 1 DPC, and platform with E5-2699 v4, Turbo enabled, 4x32GB

DDR4-2400, RHEL71, intermental measurements on platformance 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 visit http://www.intel.com/performance.

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



IARI FS SYSTFM SCAL

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.

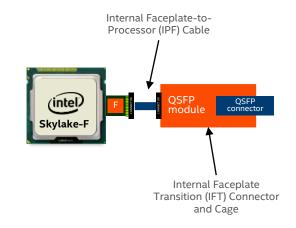
Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance



Intel[®] Xeon[®] Scalable Processor with Integrated Fabric

Single on-package Intel[®] 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 Single cable from SKL-F package connector to QSFP module Same socket for Skylake-SP and Skylake-F processors

- Intel[®] Xeon[®] Scalable 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[®] Scalable Processor Architecture Summary New Architectural Innovations for Data Center

- Up to 60% increase in compute density with Intel[®] AVX-512
- Improved performance and scalability with Mesh on-chip interconnect
- L2 and L3 cache hierarchy optimized for data center workloads
- Improved memory subsystem with up to 60% higher memory bandwidth
- Faster and more efficient Intel[®] UPI interconnect for improved scalability
- Improved integrated IO with up to 50% higher aggregate IO bandwidth
- Increased protection against kernel tampering and user data corruption
- Enhanced power management and RAS capability for improved utilization of resources



WHAT'S NEW



INTRODUCING THE NEW INTEL® XEON® SCALABLE PROCESSOR BREAKTHROUGH PERFORMANCE FOR EXPERT WORKSTATIONS® INTRODUCING THE NEW INTEL® XEON® W PROCESSOR PERFORMANCE OPTIMIZED FOR MAINSTREAM WORKSTATIONS

WTEL' KEDN' PROCESSOR V RANLY

mail Advances
 instate
 instate

(intel

intel.

INTEL' XEON' W

PROCESSOR



BREAKTHROUGH PERFORMANCE FOR EXPERT WORKSTATIONS^{*}

UP 2.71XPERFORMANCE
IMPROVEMENT
4-YEAR REFRESHUP 1.65XPERFORMANCE
IMPROVEMENT
GEN-ON-GEN2UP 56 CORESUP 112 THREADSUP 4.2 frTUB 0UP 4.2 frUP 10UP 3.0 fr266UP 3.0 fr266

ACCELERATOR THROUGHPUT WITH EXPANDABILITY, RELIABILITY, SECURITY*

INFORMATION BASED ON DUAL-SOCKET CONFIGURATION

New Intel[®] Xeon[®] Scalable Processor

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 <u>www.intel.com/benchmarks</u>. Configuration: Refer to Performance Benchmark Disclosure slide. Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance. *Other names and brands may be claimed as the property of others.

© Copyright 2017 Intel Corporation

(intel)

Intel[®] Xeon[®] Platinum Processor

WORLD RECORD PERFORMANCE



(intel) XEON

PERFORMANCE OPTIMIZED FOR MAINSTREAM WORKSTATIONS[†]

UP 1.87X PERFORMANCE TO 1.87X PERFORMANCE 4-YEAR REFRESH³ UP 1.38X PERFORMANCE TO 1.38X PERFORMANCE IMPROVEMENT GEN-ON-GEN⁴ UP 512 GB 266 \$

UP 4.5 G TURBO

TO THE CORES THE 36 THRE ADS

OPTIMIZED MAINSTREAM PERFORMANCE WITH EXPANDABILITY, RELIABILITY, SECURITY*

AVAILABLE IN SINGLE-SOCKET CONFIGURATION ONLY

New Intel[®] Xeon[®] W Processor



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 www.intel.com/benchmarks. Configuration: Refer to Performance Benchmark Disclosure slide. Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance. *Other names and brands may be claimed as the property of others.

intel

INTEL' XEON' W PROCESSOR

© Copyright 2017 Intel Corporation

NTE' JEW PROCESSIO VENUY



NEW INTEL® XEON® W PROCESSOR

Mainstream performance, enhanced memory capabilities, hardware-enhanced security and reliability features for professional workstations.

- Up to 18 cores, 36 threads
- Four channel DDR4-2666 ECC memory support
- Intel[®] Turbo Boost Technology 2.0
- Intel[®] AVX-512 acceleration with up to 2 FMA
- Support for LGA 2066 socket
- 48 PCI Express 3.0 lanes
- Intel[®] Mesh Architecture
- Intel optimized 14nm+ process technology
- Rebalanced Intel[®] smart cache hierarchy
- Intel[®] vPro[™] Technology
- Intel[®] Hyper-Threading Technology (Intel[®] HT Technology)
- Intel[®] Virtual RAID on Chip (Intel[®] VROC)
- Integrated Intel[®] Ethernet: 1 Gigabit Ethernet





