Notices and Disclaimers

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer. No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, 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.

Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel, the Intel logo, Intel Optane and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

* Other names and brands may be claimed as the property of others. © 2017 Intel Corporation.
Agenda

- Intel® Xeon® Scalable Processor Overview
- Skylake-SP CPU Architecture
- Lewisburg PCH Architecture
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 Feature Overview**

### Feature 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)

**Networking**
- Intel® Omni-Path Fabric (integrated or discrete)
  - 4x10GbE (integrated w/ chipset)
  - 100G/40G/25G discrete options

**Compression and Crypto Acceleration**
- Intel® QuickAssist Technology to support 100Gb/s comp/decomp/crypto
  - 100K RSA2K public key

**Storage**
- Integrated QuickData Technology, VMD, and NTB
  - Intel® Optane™ SSD, Intel® 3D-NAND NVMe & SATA SSD

**Security**
- CPU enhancements (MBE, PPK, MPX)
- Manageability Engine
- Intel® Platform Trust Technology
- Intel® Key Protection Technology

**Manageability**
- Innovation Engine (IE)
- Intel® Node Manager
- Intel® Datacenter Manager

---

**BMC**: Baseboard Management Controller

**PCH**: Intel® Platform Controller Hub

**IE**: Innovation Engine

**Intel® QAT**: Intel® QuickAssist Technology

**ME**: Manageability Engine

**NIC**: Network Interface Controller

**VMD**: Volume Management Device

**NTB**: Non-Transparent Bridge
INTEL® XEON® SCALABLE PROCESSOR SUPPORTS CONFIGURATIONS RANGING FROM 2S-2UPI TO 8S
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)

---

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

• Larger and improved branch predictor, higher throughput decoder, larger window to extract ILP
• 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 MLC

ABOUT 10% PERFORMANCE IMPROVEMENT PER CORE ON INTEGER APPLICATIONS AT SAME FREQUENCY
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**
  Reduced overhead for VMs moving across CPUs with 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

- **MPX (Memory Protection Extension)**
  Enables bounds checking on data accesses to prevent buffer overflow attacks

**INSTRUCTION SET ENHANCEMENT ACROSS COMPUTE, VIRTUALIZATION, AND SECURITY**
Intel® Advanced Vector Extensions 512 (Intel® AVX-512)

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

<table>
<thead>
<tr>
<th>Microarchitecture</th>
<th>Instruction Set</th>
<th>SP FLOPs / cycle</th>
<th>DP FLOPs / cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skylake</td>
<td>Intel® AVX-512 &amp; FMA</td>
<td>64</td>
<td>32</td>
</tr>
<tr>
<td>Haswell / Broadwell</td>
<td>Intel AVX2 &amp; FMA</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>Sandybridge</td>
<td>Intel AVX (256b)</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Nehalem</td>
<td>SSE (128b)</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

**Intel AVX-512 Instruction Types**

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

**POWERFUL INSTRUCTION SET FOR DATA-PARALLEL COMPUTATION**
Skylake-SP Core

Skylake-SP core builds on Skylake core with features architected for data center usage

- Intel® AVX-512 implemented with Port 0/1 fused to a single 512b execution unit
- Port 5 is extended to full 512b to add second FMA outside of Skylake core
- L1-D load and store bandwidth doubled to allow up to 2x64B load and 1x64B store
- Additional 768KB of L2 cache added outside of Skylake core
Frequency Behavior While Running Intel® AVX Code

- Cores running non-AVX, Intel® AVX2 light/heavy, and Intel® AVX-512 light/heavy code have different turbo frequency limits.
- Frequency of each core is determined independently based on workload demand.

<table>
<thead>
<tr>
<th>Code Type</th>
<th>All Core Frequency Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSE</td>
<td>Non-AVX All Core Turbo</td>
</tr>
<tr>
<td>AVX2-Light (without FP &amp; int-mul)</td>
<td></td>
</tr>
<tr>
<td>AVX2-Heavy (FP &amp; int-mul)</td>
<td>AVX2 All Core Turbo</td>
</tr>
<tr>
<td>AVX512-Light (without FP &amp; int-mul)</td>
<td></td>
</tr>
<tr>
<td>AVX512-Heavy (FP &amp; int-mul)</td>
<td>AVX512 All Core Turbo</td>
</tr>
</tbody>
</table>

Mixed Workloads

- Non-AVX_Turbo
- AVX2_Turbo
- AVX512_Turbo
- Non-AVX_Base
- AVX2_Base
- AVX512_Base
Performance and Efficiency with Intel® AVX-512

INTEL® AVX-512 DELIVERS SIGNIFICANT PERFORMANCE AND EFFICIENCY GAINS

INTEL® AVX-512 delivers significant performance and efficiency gains.

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 information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
SKYLAKE-SP SOC ARCHITECTURE
New Mesh Interconnect Architecture

Broadwell EX 24-core die

- Core
- DDR
- Cache
- SAD
- IDI/QPI
- Link
- QPI Agent
- R2PCI
- PCIe x16
- On Pkg PCIe x16
- 2.5MB LLC
- Home Agent
- Mem Ctlr

Skylake-SP 28-core die

- Core
- DDR
- Cache
- SAD
- IDI/QPI
- Link
- QPI
- R3QPI
- PCIe x16
- On Pkg PCIe x16
- 2.5MB LLC
- Home Agent
- Mem Ctlr

Mesh Improves Scalability with Higher Bandwidth and Reduced Latencies

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

• Intel® UPI caching and home agents are distributed with each LLC bank
  • Prior generation had a small number of QPI home agents
• Distributed CHA benefits
  • Eliminates large tracker structures at memory controllers, allowing more requests in flight and processes them concurrently
  • Reduces traffic on mesh by eliminating home agent to LLC interaction
  • Reduces latency by launching snoops earlier and obviates need for different snoop modes
Re-Architected L2 & L3 Cache Hierarchy

Previous Architectures

- Shared L3: 2.5MB/core (inclusive)
  - L2: (256KB private)
  - Core

Skylake-SP Architecture

- Shared L3: 1.375MB/core (non-inclusive)
  - L2: (1MB private)
  - Core

- On-chip cache balance shifted from shared-distributed (prior architectures) to private-local (Skylake architecture):
  - Shared-distributed ➔ shared-distributed L3 is primary cache
  - Private-local ➔ private L2 becomes primary cache with shared L3 used as overflow cache

- Shared L3 changed from inclusive to non-inclusive:
  - Inclusive (prior architectures) ➔ L3 has copies of all lines in L2
  - Non-inclusive (Skylake architecture) ➔ lines in L2 may not exist in L3

SKYLAKE-SP CACHE HIERARCHY ARCHITECTED SPECIFICALLY FOR DATA CENTER USE CASE
Inclusive vs Non-Inclusive L3

1. Memory reads fill directly to the L2, no longer to both the L2 and L3
2. When a L2 line needs to be removed, both modified and unmodified lines are written back
3. Data shared across cores are copied into the L3 for servicing future L2 misses

Cache hierarchy architected and optimized for data center use cases:

- Virtualized use cases get larger private L2 cache free from interference
- Multithreaded workloads can operate on larger data per thread (due to increased L2 size) and reduce uncore activity
Cache Performance

Skylake-SP cache hierarchy significantly reduces L2 misses without increasing L3 misses compared to Broadwell-EP.

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, 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. Copyright © 2017 Intel Corporation.
Cache Performance

Lower is better

Skylake-SP cache hierarchy significantly reduces L2 misses without increasing L3 misses compared to Broadwell-EP

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, 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. Copyright © 2017, Intel Corporation.
Cache Performance

CPU CACHE LATENCY

Source as of June 2017: Intel internal measurements on platform with Xeon Platinum 8180, Turbo enabled, SNC1, 6x32GB DDR4-2666 per CPU, 1 DPC, and platform with Intel® Xeon® E5-2699 v4, Turbo enabled, without COD, 4x32GB DDR4-2400, RHEL 7.0. Cache latency measurements were done using Intel® Memory Latency Checker (MLC) tool.

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. Copyright © 2017, Intel Corporation.

Skylake-SP L2 cache latency has increased by 2 cycles for a 4x larger L2

Skylake-SP achieves good L3 cache latency even with larger core count
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 XPT prefetch and D2C/D2K to reduce LLC miss latency

Introduces a new memory device failure detection and recovery scheme with Adaptive Double Device Data Correction (ADDDC)

SIGNIFICANT MEMORY BANDWIDTH AND CAPACITY IMPROVEMENTS
Sub-NUMA Cluster (SNC)

Prior generation supported Cluster-On-Die (COD)

SNC provides similar localization benefits as COD, without some of its downsides

- Only one UPI caching agent required even in 2-SNC mode
- Latency for memory accesses in remote cluster is smaller, no UPI flow
- LLC capacity is utilized more efficiently in 2-cluster mode, no duplication of lines in LLC
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

Higher is better

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® 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

**Data Rate**

<table>
<thead>
<tr>
<th></th>
<th>QPI</th>
<th>UPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Rate</td>
<td>9.6 GT/s</td>
<td>10.4 GT/s</td>
</tr>
</tbody>
</table>

**Data Efficiency**

<table>
<thead>
<tr>
<th></th>
<th>LO QPI</th>
<th>L0p UPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Efficiency</td>
<td>4% to 21%</td>
<td>75% to 50%</td>
</tr>
</tbody>
</table>

**Idle Power**

<table>
<thead>
<tr>
<th></th>
<th>LO QPI</th>
<th>L0p UPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle Power</td>
<td>75%</td>
<td>50%</td>
</tr>
</tbody>
</table>

**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](http://www.intel.com/performance).
Processor Integrated I/O

3 independent pipelines of x16 PCIe* Gen3
- Each x16 can be bifurcated into 2x8, 1x8+2x4, or 4x4 root ports
- New traffic controller pipeline improves over prior design
- Additional x16 PCIe for Intel® Omni-Path integration

Non-Transparent Bridging (NTB)
- One NTB per x16 PCIe, which can be configured as 1x8 or 1x4 NTB

Intel® QuickData Technology (CBDMA)
- 2x bandwidth on Mem-Mem copy
- Supports MMIO-Mem copy

Intel® Volume Management Device (VMD)
- One VMD domain per x16 PCIe

MODULAR IO DESIGN WITH IMPROVED FEATURE SET FOR CONVERGED DATA CENTER
## IO Performance

### IO Bandwidth Change Over Xeon E5-2699V4

<table>
<thead>
<tr>
<th>Relative IO Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEMORY BW (READ)</td>
</tr>
<tr>
<td>AGG. LOCAL PCIe</td>
</tr>
<tr>
<td>RDWR (LARGE PKT)</td>
</tr>
<tr>
<td>DMA MEM-MEM</td>
</tr>
<tr>
<td>DMA MEM-IO</td>
</tr>
<tr>
<td>DMI</td>
</tr>
</tbody>
</table>

| Source as of June 2017: Intel internal measurements on platform with Xeon Platinum 8180, Turbo enabled, UPI=10.4, SNC1, 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 complete information visit [http://www.intel.com/performance](http://www.intel.com/performance).
SKYLAKE-SP NEW CAPABILITIES
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**
Energy Efficiency and Power Management Enhancements

**ENERGY EFFICIENCY**

- Optimized intermediate core turbo profile
- Dynamic power sharing between core, uncore, and fabric HFI
- Larger L2 cache for reduced interconnect and coherency activity
- Power delivery through integrated VR for core and uncore

**POWER MANAGEMENT**

- Intel® Speed Shift Technology for autonomous P-state control
- Improved core and uncore frequency scaling heuristics
- On-die Pmax detector for rapid response to power excursions
- Independent per core and CLM (CHA, LLC, and mesh) voltage and frequency domains
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.
Intel® Speed Shift Technology Interface

Legacy Interface

- P0 – 1 core
- P0 – 2 cores
- P1
- P2
- Pn
- T-states

Turbo frequency
Guaranteed frequency
Energy efficient Frequency (min V)
Thermal control

DVFS – Intel SpeedStep® Technology
\[ P \sim V^2 \cdot f \cdot C_{dyn} + \text{leakage}(V) \sim f^3 \]

Legacy: OS controls P-state
- P1-Pn enumerated via ACPI tables
- Explicit P-state selection to P1
- Autonomous control in turbo range

New: guided autonomous control
- OS provides min, max and preference
- Demand based HW control

HWP i/F

- P0 – 1 core
- Maximum frequency
- Optional: Desired frequency
- Minimum frequency
- Lowest frequency
- LFM
- T-states
**Intel® Xeon® Scalable Processor:**
Intel® Run Sure Technology Features

### Resilient System Technologies
- Advanced Error Detection and Correction (AEDC)
- MCA 2.0 Recovery (as per eMCA gen2 architecture)
- MCA Recovery-Execution Path
- MCA Recovery-Non Execution Path
- Local Machine Check (LMCE) based recovery

### Resilient Memory Technologies
- SDDC + 1, Adaptive DDDC (MR) +1
- Addressed Range/Partial Memory Mirroring

**Resilient System Technologies** - integrate processor, firmware, and software layers that allow the system to diagnose and/or recover from previously fatal errors

**Resilient Memory Technologies** - ensure data integrity & enable systems to keep running reliably over a longer period of time, reducing the frequency of service calls

**CONTINUED IMPROVEMENT IN DATA CENTER UPTIME WITH INTEL® RUN SURE TECHNOLOGY**
Intel® Xeon® Scalable Processor: New Virtualization Enhancements

GREATER CONSOLIDATION ON A COMPUTE NODE
- Improved core performance and larger number of cores
- Larger TLBs and per core L2 cache for improved performance and lower variability on virtualized workloads
- Improved memory bandwidth and capacity to collocate demanding workloads

SECURE CREATION AND MOVEMENT OF VM ACROSS SYSTEMS
- Mode Based Execution Control for hardening VM launch vulnerability
- Improved timestamp virtualization to reduce overhead of migrating VMs across different CPU skus
Intel® Xeon® Scalable Processor: New Security Enhancements

**SECURITY PERFORMANCE**
- Intel® AVX-512
- Intel® QuickAssist Technology

**HARDENING AGAINST ATTACK SURFACES**
- Mode Based Execution to protect from attacks during VM creation
- Page Protection Keys to create robust applications with differentiated page access rights managed at user level
- Intel® Memory Protection Extension (Intel® MPX) to prevent buffer overflow attacks

**HARDENING THE PLATFORM**
- Secure key management within the chipset without a discrete TPM
- Authenticated and measured launch and recovery options
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
SKYLAKE-SP CPU WRAP UP
Skylake-SP 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
- Core, cache, memory and IO improvements for **increased virtual machine performance**
- **Enhanced power management and RAS capability** for improved utilization of resources
Skylake-SP Performance

Skylake-SP CPUs provide significant performance upside compared to prior generation

165W Skylake-SP CPUs provide more than 40% gain on performance

205W Skylake-SP CPUs provide additional boost to core bound workloads

Source as of June 2017: Intel internal measurements with Xeon Platinum 8180 and 8176, Turbo enabled, UPI=10.4, SNC1, 6x32GB DDR4-2666 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 the product when combined with other products. For more complete information visit http://www.intel.com/performance.
Skylake-SP Die Configurations

XCC Die with 28 Cores

- 6x6 Mesh topology
- 5 rows of core and LLC
- 2 memory controllers, one on each side of die
- All IOs at the top
- 3 x16 PCIe Gen3 stacks
- 1 x16 PCIe for MCP use
- Up to 3 Intel® UPI ports
High and Low Core Count Die Configurations

**HCC (up to 18 cores)**

- 2x UPI x 20 @ 10.4GT/s
- 1x16/2x8/4x4 PCIe @ 8GT/s x4 DMI
- 1x16/2x8/4x4 PCIe @ 8GT/s x4 DMI
- 2x PCIe x 16
- 2x PCIe x 16
- 2x PCIe x 16
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4

**LCC (up to 10 cores)**

- 2x UPI x 20 @ 10.4GT/s
- 1x16/2x8/4x4 PCIe @ 8GT/s x4 DMI
- 1x16/2x8/4x4 PCIe @ 8GT/s x4 DMI
- 2x PCIe x 16
- 2x PCIe x 16
- 2x PCIe x 16
- 2x PCIe x 16
- 2x PCIe x 16
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x CHA/SF/LLC
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4
- 2x DDR4

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

*This slide under embargo until 9:15 AM PDT July 11, 2017*
LEWISBURG CHIPSET ARCHITECTURE
Lewisburg: New PCH for the Converged Platform

Ubiquity of network security, efficient data storage and packet manipulation in Cloud, Storage, Enterprise and Network appliances.

Lewisburg is a Common Platform Offering in Purley generation to meet the converged requirement

Data Center PCH provides boot, standard legacy and high-speed IO, manageability and clocking solutions

- Intel® Innovation Engine is a manageability sandbox for system builder

Integrated Intel® Ethernet Controller reduces area, power, cost and provides platform LOM capabilities

- Intel® Ethernet Connection X722 with up to 4x10Gbps
- Supports Network Virtualization Offloads and iWARP RDMA

Intel® QuickAssist Technology Accelerators provide security and compression for Communications, Storage, and Cloud deployments

- Crypto / compression up to 100Gbps
# PCH Generational Comparison

<table>
<thead>
<tr>
<th>Features</th>
<th>Intel® C610 series chipset (Wellsburg)</th>
<th>Intel® C620 series chipset (Lewisburg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Features</td>
<td>Intel® vPro™ technology, AMT, Node Manager 3.0; Server Trace Lengths; 6 SMBus Discrete, Integrated and Hybrid Clocking</td>
<td>Intel® vPro™, AMT, Node Manager 4.0; Server Trace Lengths; 6 SMBus Discrete, Integrated and Hybrid Clocking</td>
</tr>
<tr>
<td>SATA Ports</td>
<td>Up to 10 SATA 3 (6 Gb/s)</td>
<td>Up to 14 SATA 3 (6 Gb/s)</td>
</tr>
<tr>
<td>USB Ports</td>
<td>Up to 14 USB 2.0; Up to 6 USB 3.0</td>
<td>Up to 14 USB 2.0; Up to 10 USB 3.0</td>
</tr>
<tr>
<td>DMI</td>
<td>x4, 2.0 speed</td>
<td>x4, 3.0 speed</td>
</tr>
<tr>
<td>Additional CPU Uplink Options</td>
<td>N/A</td>
<td>PCIe® 3.0 at x8 and/or x16;</td>
</tr>
<tr>
<td>PCI Express*</td>
<td>Up to 8 PCIe 2.0 (5 GT/s)</td>
<td>Up to 20 ports PCIe 3.0 (8 GT/s)</td>
</tr>
<tr>
<td>TPM Support</td>
<td>TPM 1.2</td>
<td>TPM 2.0</td>
</tr>
<tr>
<td>SATA Express, NVM Express</td>
<td>No</td>
<td>No / Supported</td>
</tr>
<tr>
<td>Intel® RSTe</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Innovation Engine</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>LAN</td>
<td>Integrated 1GbE</td>
<td>Integrated Intel® Ethernet Connection X722 with up to 4x10Gb/1Gb ports with iWARP RDMA/ SFI</td>
</tr>
<tr>
<td>Intel® QuickAssist Technology - Crypto</td>
<td>N/A</td>
<td>Up to 100 Gb/s IPSec/SSL</td>
</tr>
<tr>
<td>Intel QuickAssist Technology - Public Key</td>
<td>N/A</td>
<td>Up to RSA2K 100K Ops</td>
</tr>
<tr>
<td>Intel QuickAssist Technology – Compression</td>
<td>N/A</td>
<td>Up to 100Gb/s (Deflate(LZ77))</td>
</tr>
<tr>
<td>Enhanced Serial Peripheral Interface (eSPI)</td>
<td>N/A</td>
<td>up to 60MHz, 1.8V</td>
</tr>
</tbody>
</table>

**INTEGRATED 10Gb INTEL® ETHERNET, ENHANCED I/O AND PLATFORM SECURITY WITH ROBUST CRYPTO AND COMPRESSION SOLUTIONS**

*Green text is new*
Lewisburg Deployment Configurations

**Conventional**
On-board PCH. Dual PCIe uplink to single CPU, provisioned for Intel® QAT and 4x10GbE (PCIe Uplink optional)

**CPU Straddling**
On-board PCH PCIe uplink to multiple CPU, provisioned for QAT and 4x10GbE

**PCIe* Endpoint Mode**
On-board or PCIe-card-based acceleration for scalability, provisioned for QAT and 4x10GbE

**Wrap-Around**
On-board PCH Configure low-BW QAT/Ethernet as PCH-attached downstream device

**Multi-PCH**
Multiple On-board PCH for failover, flexible partition or IO expansion
Optional PCIe uplink for QAT and 4x10GbE
Intel® QuickAssist Technology

Generation 2

- **Secure Key Establishment**
  - AES256 + HMAC SHA256 (150Gbs)
  - RSA 2048 Decrypt (100kops)
- **Lossless compression for data in flight & rest**
  - Deflate Compression (100+Gbs)

**Bulk Crypto**
- Security for data in flight & rest

**Public Key Encryption**
- Secure Key Establishment

**Compression**
- Lossless compression for data in flight & rest

INTEL® QUICKASSIST TECHNOLOGY IS DESIGNED TO OPTIMIZE THE USE & DEPLOYMENT OF CRYPTO AND COMPRESSION HARDWARE ACCELERATORS ON INTEL® PLATFORMS
Intel® QuickAssist Technology

Generation 2

**Bulk Crypto**
- **300%**
- **150Gbs**
  - AES256 + HMAC SHA256

**Public Key Encryption**
- **250%**
- **100kops**
  - RSA 2048 Decrypt

**Compression**
- **400%**
- **100+Gbs**
  - Deflate Compression

**Security for data in flight & rest**

**Secure Key Establishment**

**Lossless compression for data in flight & rest**

**New Platform Capability**

**Intel® Key Protection Technology**

---

**INTEL® QUICKASSIST TECHNOLOGY IS DESIGNED TO OPTIMIZE THE USE & DEPLOYMENT OF CRYPTO AND COMPRESSION HARDWARE ACCELERATORS ON INTEL® PLATFORMS**

*Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests 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/datacenter](http://www.intel.com/performance/datacenter). Configurations: see slide 71.*
Intel® QuickAssist Technology: Crypto

Usage Model
- Network security (IPsec, SSL/TLS), hashing for data-deduplication, encrypted storage

Symmetric (Bulk) Cryptography
- Ciphers (AES, 3DES/DES, RC4, KASUMI®, Snow 3G)
- Message digest/hash (MD5, SHA1, SHA2x, SHA3) and authentication (HMAC, AES-XCBC)
- Algorithm chaining (one cipher and one hash in a single operation)
- Authenticated encryption (AES-GCM, AES-CCM)
- AES-XTS

Wireless
- KASUMI, Snow 3G and ZUC

Asymmetric (Public Key) Cryptography
- Modular exponentiation for Diffie–Hellman (DH)
- RSA key generation, encryption/decryption and digital signature generation/verification
- DSA parameter generation and digital signature generation/verification
- Elliptic Curve Cryptography: ECDSA, ECDHE

### Performance

<table>
<thead>
<tr>
<th>Network Security Protocols</th>
<th>Coleto Creek</th>
<th>Lewisburg</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLS @ 16k records</td>
<td>50 Gbs</td>
<td>150 Gbs¹</td>
</tr>
<tr>
<td>IPSec @ 1kB</td>
<td>45 Gbs</td>
<td>100 Gbs¹</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Public Key Encryption</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RSA Decrypt 2K</td>
<td>40k Ops</td>
<td>100k Ops</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wireless Ciphers</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ZUC/Snow 3G/KASUMI® F8²</td>
<td>20 Gbs</td>
<td>50Gbs</td>
</tr>
</tbody>
</table>

Ciphers or Hash Only

| AES128 CBC @ 4k           | 50Gbs         | 150Gbs    |
| SHA1, SHA256, SHA3, MD5 @ 4k | 50Gbs        | 140Gbs    |

¹ Using 16k records using AES-CBC-HMAC SHA1/256, 100Gbs at 1k packet
² KASUMI-F8 (encryption) at 320B packets, 7Gbs for 64B packets

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.

Intel measurements as of May 2017: Intel® Customer Reference Board Neon City with one Intel® Xeon® Skylake-SP processor and one Lewisburg Chipset LBG-T 80 ES2* SKU (DCL: #568586). For more complete information about performance and benchmark results, visit www.intel.com/benchmarks
Intel® QuickAssist Technology: Data Compression

Usage Model
- Big data acceleration
- WAN acceleration
- Http compression
- File System
- Databases

Compression and Decompression Algorithm
- DEFLATE: LZ77 compression followed by Huffman coding, with a gzip or zlib header

Other Features
- Engine can be configured to perform either compression or decompression
- Support for stateful (de)compression
- Storage-specific features

Performance

<table>
<thead>
<tr>
<th></th>
<th>Coleta Creek</th>
<th>Lewisburg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compression</td>
<td>24 Gbs(^1)</td>
<td>100+ Gbs(^1)</td>
</tr>
<tr>
<td>Decompression</td>
<td>24 Gbs(^1)</td>
<td>160 Gbs(^1)</td>
</tr>
<tr>
<td>Compression + Decompression</td>
<td>24 Gbs(^1)</td>
<td>100 Gbs(^1)</td>
</tr>
</tbody>
</table>

\(^1\) Dynamic Deflate Level 1 using 64KB buffer size
\(^2\) Best case compression ratio with Lewisburg is zlib level 4

Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. 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.

Intel measurements as of May 2017; Intel® Customer Reference Board Neon City with one Intel® Xeon® Skylake-SP processor and one Lewisburg Chipset LBG-T B0 ES2\(^a\) SKU (DCL: #568586). For more complete information about performance and benchmark results, visit [www.intel.com/benchmarks](http://www.intel.com/benchmarks)
Integrated Intel® Ethernet Connection X722

Low-cost solution for up to four ports of 10Gb Ethernet
- >50% of server ports will be 10GbE in 2017\(^1\)
- Lower power, less board space: integrated 10GbE saves > 20 cm\(^2\) in board area and > 25% in power consumption \(^2\)
- Available as integrated Ethernet on motherboard or stand-alone NIC cards

Proven “It Just Works” 10GbE solution
- Based on Intel® Ethernet Converged Network Adapter XL710 (Fortville) IP
- External PHYs in production today support 1GbE and 10GbE 10GBASE-T

Advanced features to enable Software Defined Infrastructure
- Network Virtualization Offloads to support the move to L3 networks
- iWARP RDMA for increased bandwidth at lower CPU utilization
- Intel® Ethernet Flow Director traffic steering for increased efficiency
- Intel® Data Plane Development Kit (DPDK) for advanced packet forwarding

---

# Intel® 4x10Gb/1Gb Ethernet Controller

<table>
<thead>
<tr>
<th>Feature Support</th>
<th>Customer Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quad Port 10GbE MAC/PHY</td>
<td>• Optimized for networking capability in Cloud, Comms, and Storage</td>
</tr>
<tr>
<td>• Based on Intel's 10GbE solution</td>
<td>• Single network driver on Intel® platform</td>
</tr>
<tr>
<td>• Interfaces: 10G (KR, SFI, XFI), 1G (KX)</td>
<td></td>
</tr>
<tr>
<td>Remote Direct Memory Access</td>
<td>• Routable and scalable RDMA ideal for large segmented networks in private and public clouds</td>
</tr>
<tr>
<td>• iWARP</td>
<td></td>
</tr>
<tr>
<td>Network Virtualization Offloads</td>
<td>• Abstract the network for cloud flexibility</td>
</tr>
<tr>
<td>• Flexible Filters (ATR, Flow Director)</td>
<td>• Enhanced programmability and application affinity</td>
</tr>
<tr>
<td>• NVGRE, IPinGRE, VXLAN, MACinUDP</td>
<td></td>
</tr>
<tr>
<td>Standards Based Virtualization</td>
<td>• Broad OS enablement</td>
</tr>
<tr>
<td>• SR-IOV: 4 Physical/128 Virtual Function</td>
<td></td>
</tr>
<tr>
<td>• VEB (Virtual Ethernet Bridge)</td>
<td></td>
</tr>
<tr>
<td>Power Management</td>
<td>• Energy Efficient Ethernet</td>
</tr>
<tr>
<td>Manageability</td>
<td>• Common transport for LAN-to-BMC</td>
</tr>
<tr>
<td>• BMC Pass-Through (control &amp; network)</td>
<td></td>
</tr>
<tr>
<td>• Interfaces: NC-SI, SMBus, and MCTP</td>
<td></td>
</tr>
</tbody>
</table>

## Feature Support

- **Quad Port 10GbE MAC/PHY**
  - Based on Intel's 10GbE solution
  - Interfaces: 10G (KR, SFI, XFI), 1G (KX)

- **Remote Direct Memory Access**
  - iWARP

- **Network Virtualization Offloads**
  - Flexible Filters (ATR, Flow Director)
  - NVGRE, IPinGRE, VXLAN, MACinUDP

- **Standards Based Virtualization**
  - SR-IOV: 4 Physical/128 Virtual Function
  - VEB (Virtual Ethernet Bridge)

- **Power Management**

- **Manageability**
  - BMC Pass-Through (control & network)
  - Interfaces: NC-SI, SMBus, and MCTP

## Customer Value

- **Quad Port 10GbE MAC/PHY**
  - Optimized for networking capability in Cloud, Comms, and Storage
  - Single network driver on Intel® platform

- **Remote Direct Memory Access**
  - Routable and scalable RDMA ideal for large segmented networks in private and public clouds

- **Network Virtualization Offloads**
  - Abstract the network for cloud flexibility
  - Enhanced programmability and application affinity

- **Standards Based Virtualization**
  - Broad OS enablement

- **Power Management**
  - Energy Efficient Ethernet

- **Manageability**
  - Common transport for LAN-to-BMC
Remote Direct Memory Access (RDMA)

RDMA is a network performance optimization
• Enables direct app-to-app communication across nodes
• Bypasses the OS stack & kernel
• Provides direct channel for remote memory application access

RDMA offers lower latency, high throughput by:
• Avoiding application context switching
• Placing data directly in application buffers (zero-copy DMA)
• Moving protocol processing off the CPU

Intel® Ethernet Connection X722 is featured with iWARP (Internet Wide-area RDMA Protocol) RDMA:
• Recommended for easy deployment and configuration, scalability and congestion control

RDMA-aware applications or messaging layers are required
• Windows*/Linux* Drivers supported for Storage, Messaging Middleware and HPC application categories
Introducing the Innovation Engine (IE)

Embedded core in the LBG PCH
- Very small Intel® Architecture core
- Similar to Management Engine (ME) hardware, with some privilege and IO differences

Reserved for system-builder's code, not for Intel firmware
- Intel supplies IE hardware only
- IE code is cryptographically bound to the system-builder
- Code not authenticated by the system-builder will not load

Activation is not required for normal system operation
- Optional feature
- Requires SPS firmware. ME1x firmware does not support

*BMC is optional if system-builder has implemented a management app in IE that communicates directly with the network.
## Lewisburg ME vs IE Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>LBG ME</th>
<th>LBG IE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Intel® Quark™ x86 (412 DMIPS)</td>
<td>Intel Quark x86 (412 DMIPS)</td>
</tr>
<tr>
<td>Memory</td>
<td>1.7MB SRAM</td>
<td>1.4MB SRAM</td>
</tr>
<tr>
<td>Crypto Algorithms</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Host Interface</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Platform Interface</td>
<td>Yes</td>
<td>Yes + UART</td>
</tr>
<tr>
<td>FW</td>
<td>Intel® AMT/vPro, SPS</td>
<td>System Builder FW</td>
</tr>
<tr>
<td>Security</td>
<td>TXT, BtG, PTT, KPT</td>
<td>-</td>
</tr>
<tr>
<td>Operating States</td>
<td>S0/M0, Sx/M3, Sx-Moff</td>
<td>S0/I0, Sx/I3, S0/Ioff, Sx/Ioff</td>
</tr>
</tbody>
</table>

Red: delta between ME and IE
Broad IE Usage Models

**Lite, BMC-less Manageability**

**Segment**  
Scale-out Cloud & Embedded/Appliance

**Value**  
Basic manageability without cost, space, power of a BMC  
Common manageability solution across vendors

**Usages**  
Simplified platform management (IPMI, Redfish, SoftKVM) without a BMC  
Hardware Security Module  
Basic node management

**BMC/BIOS/ME Assist**

**Segment**  
Enterprise & Scale-up Cloud

**Value**  
Improved system manageability through tighter platform integration  
Better performance through reduced BMC/CPU interrupts

**Usages**  
In-band PECI over DMI link: higher bandwidth, lower latency internal data paths  
Offload tight-loop sensor monitoring  
Power sequencing & control  
Error & reset handling  
OOB telemetry for performance and predictive RAS  
Platform NVRAM firmware security (validate BMC/etc)  
Crash dump & recovery over PECI  
Platform OOB key mgmt
## Lewisburg PCH SKU Guidance

<table>
<thead>
<tr>
<th>Product Name</th>
<th>SKU</th>
<th>10Gb/1Gb Ethernet Ports*</th>
<th>Compression</th>
<th>Encryption</th>
<th>RSA</th>
<th>Max PCIe* Uplink</th>
<th>Recommended Min Uplink Config</th>
<th>PCIe* Uplink x8 Optional Muxed Link</th>
<th>Est TDP (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intel® C621 Chipset</td>
<td>LBG-1G</td>
<td>0/4</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>x1</td>
<td>x1</td>
<td>n/a</td>
<td>~ 15</td>
</tr>
<tr>
<td>Intel® C622 Chipset</td>
<td>LBG-2</td>
<td>2/4†</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>x8</td>
<td>x4</td>
<td>n/a</td>
<td>~ 17</td>
</tr>
<tr>
<td>Intel® C624 Chipset</td>
<td>LBG-4</td>
<td>4/4</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>x16</td>
<td>x8</td>
<td>n/a</td>
<td>~ 19</td>
</tr>
<tr>
<td>Intel® C625 Chipset</td>
<td>LBG-E</td>
<td>4/4</td>
<td>20 Gb/s</td>
<td>20 Gbs</td>
<td>20K Ops</td>
<td>x16</td>
<td>x16</td>
<td>n/a</td>
<td>~ 21</td>
</tr>
<tr>
<td>Intel® C626 Chipset</td>
<td>LBG-M</td>
<td>4/4</td>
<td>40 Gb/s</td>
<td>40 Gbs</td>
<td>40K Ops</td>
<td>x16</td>
<td>x16</td>
<td>enabled</td>
<td>~ 23</td>
</tr>
<tr>
<td>Intel® C627 Chipset</td>
<td>LBG-T</td>
<td>4/4</td>
<td>100 Gb/s</td>
<td>100 Gbs</td>
<td>100K Ops</td>
<td>x16</td>
<td>x16</td>
<td>enabled</td>
<td>~ 26</td>
</tr>
<tr>
<td>Intel® C628 Chipset</td>
<td>LBG-L</td>
<td>4/4</td>
<td>100 Gb/s</td>
<td>100 Gbs</td>
<td>100K Ops</td>
<td>x16</td>
<td>x16</td>
<td>enabled</td>
<td>~ 21</td>
</tr>
</tbody>
</table>

---

*Four ports total: ports 0 & 1 can run up to 10 GbE, while ports 2 & 3 are limited to 1 GbE

Package Size (all SKUs): 34x28mm, Package Pin Count: 1310

Intel recommends two lanes of PCIe3 for each active 10GbE port for networking and x16 if Intel® QuickAssist Technology is active. LBG supports x16, x8, x4 and x1 options, up to the maximum uplink width.

†These Ethernet ports are in addition to the 1Gb port used by ME11.6.

All SKUs, frequencies and features are PRELIMINARY and can change without notice.
Summary of LBG HSIO Features

**DMI**
- Dedicated Gen3 x4, 4GB/s raw bidirectional throughput

**PCIe* Uplink**
- Dedicated Gen3 1 x16 controller
- Configurable Gen3 1 x8 controller

**PCIe* Downlink**
- Configurable Gen3, 5 x4 controllers, total of 20 lanes

**SATA**
- Configurable Gen3, 14 ports over 2 controllers
- Soft strap or GPIO based selection of PCIe or SATA

**USB**
- Configurable 10 USB3.0 lanes, 14 USB2.0 lanes
- Closed-Chassis Debug (DCI)

**GbE:**
- 1 MAC port, configurable over 5 lanes

---

**Flex IO Lane Muxing**

<table>
<thead>
<tr>
<th>Lane #</th>
<th>PCIe Down</th>
<th>PCIe Up</th>
<th>SATA</th>
<th>USB</th>
<th>GbE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td>3.0</td>
</tr>
<tr>
<td>8</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td>6x1</td>
</tr>
<tr>
<td>12</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>PCIe x4</td>
<td>PCIe x2</td>
<td>x1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This slide under embargo until 9:15 AM PDT July 11, 2017
Other Features

**eSPI (Enhanced Serial Peripheral Interface):**
- High-speed, low pin count LPC replacement interface
- Simplified routing, low-cost BMC enabling and BOM reduction
- Supports PCH or BMC-attached Flash sharing, GPIO virtualization

**Power management:**
- ASPM S-states and DeepS5 support. 2b VID control of logic core voltage
- Independent host, CSME, and IE operating states

**RAS:**
- ADR: Asynchronous DRAM Self-Refresh
- DWR: Demoted Warm Reset
INTEL® XEON® SCALABLE PROCESSOR

The Secure, Agile, Next-Generation Platform for Multi-Cloud Infrastructures

Pervasive Performance for Actionable Insights

Skylake-SP cores
Intel® AVX-512
Feeds: UPI, 6x DDR4, 3x16 PCIe,
Intel® SSDs
Integration: Intel® Ethernet / Omni-Path / Intel® QuickAssist / FPGA

Security Without Compromise

Intel® AVX-512
PPK, MPX, MBE
Intel® QAT w/ Secure Key Management
Intel® Trusted Infrastructure
Intel® Boot Guard

Agile Service Delivery

Intel® Volume Management Device Technology
Intel® RAS
Open Stack Software Optimizations
# New Memory Latency and Bandwidth Optimizations

## CACHE & MEMORY

- **Sub-NUMA Cluster Mode**
  Associates LLC slice with nearest memory controller
  Applications use NUMA primitives to achieve lower LLC/memory latency

- **XPT Prefetch**
  Core miss initiates local memory access in parallel with LLC access
  Uses history-based prediction to avoid unnecessary prefetches

- **Local and Remote Direct-to-Core**
  Data sent directly from memory controller or UPI to requesting core

## INTEL® UPI

- **UPI Prefetch**
  Similar to XPT prefetch, but for UPI requests from remote socket

- **Direct-to-UPI**
  For UPI requests from remote socket, memory controller directly sends data to UPI instead of going through CHA

- **UPI Optimizations**
  - HitME cache – directory cache for frequently used lines
  - IO Directory Cache – directory cache for remote IO writes
  - Opportunistic Snoop Broadcast (OSB) – Avoids directory lookup and update for local InvItoE
Sub-NUMA Clusters – 2 SNC Example

SNC partitions the LLC banks and associates them with memory controller to localize LLC miss traffic

- LLC miss latency to local cluster is smaller
- Mesh traffic is localized, reducing uncore power and sustaining higher BW

Without SNC

Local SNC Access
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® Speed Shift Technology

Hardware P-state (HWP) is a new capability for cooperative hardware + software performance control

• Hardware monitors activity / scalability and selects frequency at much faster time scale

• Node Manager or OS provides guidance and constraints to direct hardware
  – Minimum Performance Level: A performance floor for meeting QoS requirements
  – Maximum Performance Level: A ceiling to limit low priority applications and services

Node Manager or OS is no longer required to monitor activity and update frequency requests at regular intervals

• Node Manager or OS can update HWP guidance based on important events
  – Change in administrator settings, real-time process started, etc.
Intel® Speed Shift Technology

What is the Energy Performance Preference?

- Specifies software preference towards high performance, low power, or some balance between the two

![Graph showing the impact of Energy Performance Preference on power and performance](image)
Lewisburg - New Chipset for the Converged Platform

Provides a common footprint from entry level to full configuration with integrated classic server functions

LAN: Intel® Ethernet Connection X722 with up to 4x10GbE

- Supports Network Virtualization Offloads and iWARP RDMA

Comms and Storage accelerator: Intel® QuickAssist Technology (Intel® QAT)

- Crypto / compression up to 100Gbps
- Public key exchange up to 100K Ops

*Other names and brands may be claimed as the property of others.
Lewisburg PCH: Intel® QuickAssist Technology

- Enables standard server platforms to offer ubiquitous compression and security features
- Networking, Storage, Big Data, Cloud, HPC, and Data center applications achieve high performance on:
  - Bulk Ciphers, Authentication
  - Public Key Cryptography
  - Compression

![Graph showing performance comparison between 8920 (Cave Creek), 8955 (Coleto Creek), and Lewisburg]
## Configuration: Intel® QAT

<table>
<thead>
<tr>
<th>QAT API Level tests</th>
<th>1-Node, 1 x Intel® Xeon® Platinum 8180 Processor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Intel(R) Xeon(R) Gold 6152 H0 (30M Cache, 2.1 GHz)</td>
</tr>
<tr>
<td>Vendor</td>
<td>Intel</td>
</tr>
<tr>
<td>Nodes</td>
<td>1</td>
</tr>
<tr>
<td>Sockets</td>
<td>1</td>
</tr>
<tr>
<td>Cores Per Processor</td>
<td>22</td>
</tr>
<tr>
<td>Logical Processors</td>
<td>44</td>
</tr>
<tr>
<td>Platform</td>
<td>Purley-EP (Lewisburg – B1 stepping)</td>
</tr>
<tr>
<td>Accelerator Used</td>
<td>Intel Lewisburg in x24 link mode</td>
</tr>
<tr>
<td>Platform Comments</td>
<td>Neon City</td>
</tr>
<tr>
<td>Memory DIMMs Slots used/Processor</td>
<td>6</td>
</tr>
<tr>
<td>Total Memory</td>
<td>96 GB</td>
</tr>
<tr>
<td>Memory DIMM Configuration</td>
<td>16 GB / 2666 MT/s / DDR4 RDIMM</td>
</tr>
<tr>
<td>Memory Comments</td>
<td>Kingston 9965662-009.A00G, 16GB, 2Rx8</td>
</tr>
<tr>
<td>Network Interface Cards</td>
<td>3x XL710 X710 (Quad Fortville Card) , 12 10Gb ports used</td>
</tr>
<tr>
<td>OS</td>
<td>Ubuntu 16.10</td>
</tr>
<tr>
<td>OS/Kernel Comments</td>
<td>4.8</td>
</tr>
<tr>
<td>Primary / Secondary Software</td>
<td>QAT1.7.Upstream.L.1.0.0-15</td>
</tr>
<tr>
<td>Computer Type</td>
<td>Server</td>
</tr>
<tr>
<td>Benchmark</td>
<td>QAT API Level Sample tests</td>
</tr>
</tbody>
</table>