

Neue Dual-CPU Server mit Intel Xeon Scalable Performance (Codename Purley/Skylake-SP)



 @wefinet

Werner Fischer, Thomas-Krenn.AG

Webinar, 17. Oktober 2017

TH=MAS
KRENN®

Intel Xeon Scalable Performance

- Das ist NEU:
 - Neue Mikroarchitektur
 - AVX512
 - Größere L2-Caches
 - Memory Subsystem – 6 Channels/CPU
 - RAS Funktionen
- Die Prozessoren: Xeon Bronze, Silver, Gold & Platinum
- Skylake-SP Server bei Thomas-Krenn

Eine CPU für 1, 2, 4 und 8 Sockets



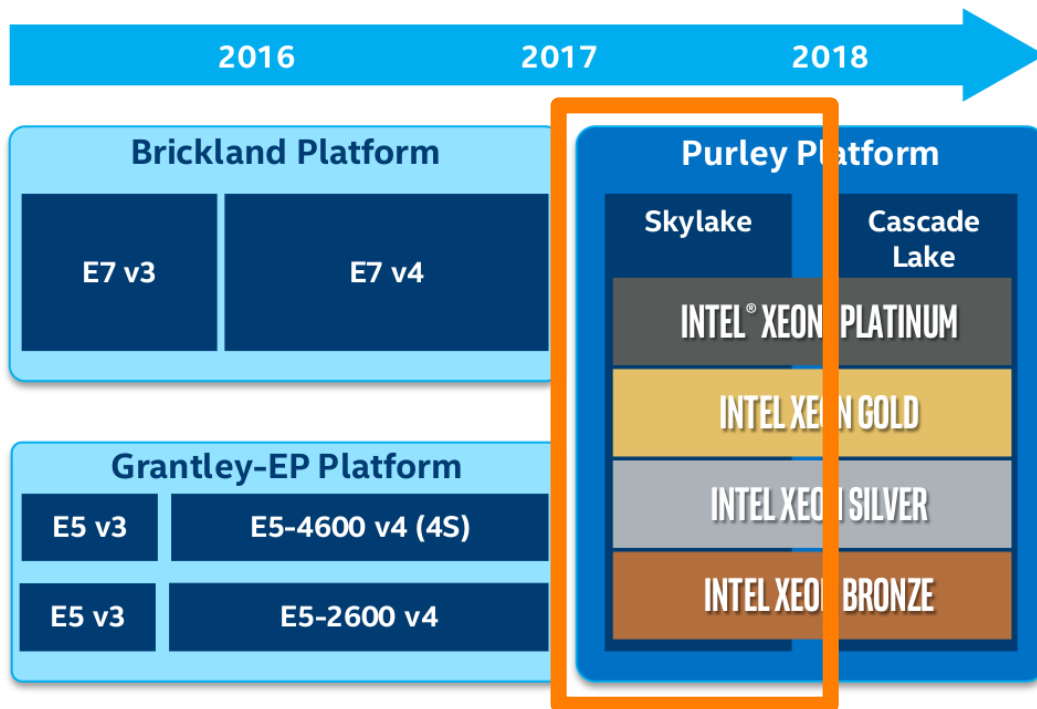
Intel® Xeon® Processor E7

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

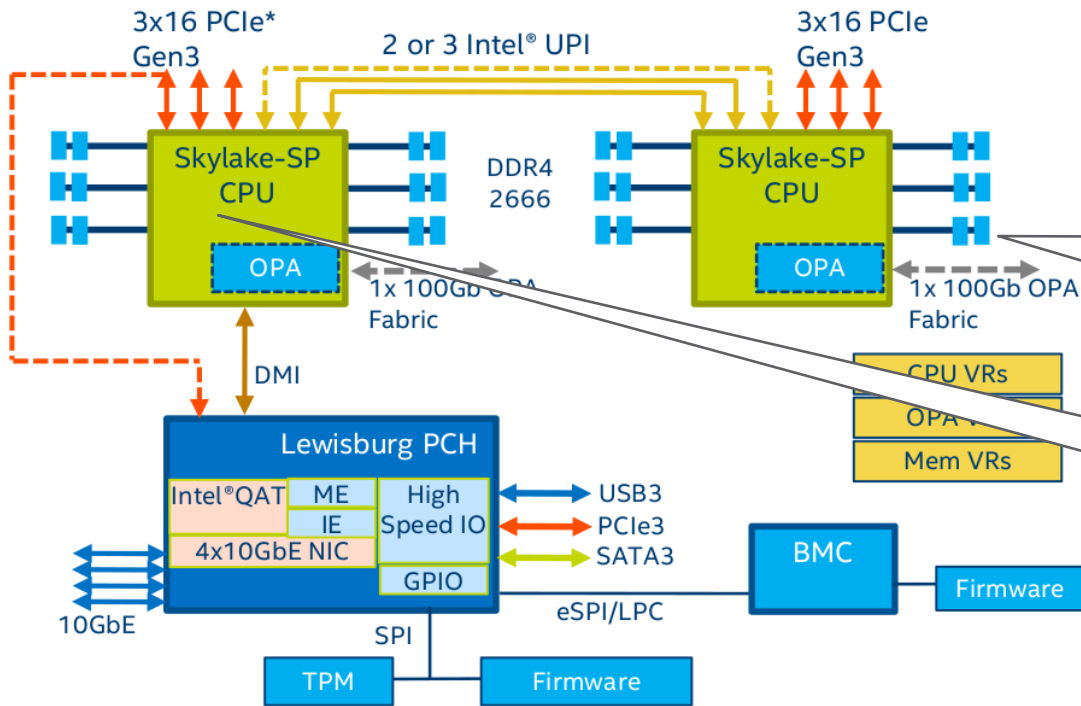


Intel® Xeon® Processor E5

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



Neue Mikroarchitektur



48 (3x16) PCIe 3.0 Lanes
je CPU (bisher 44)

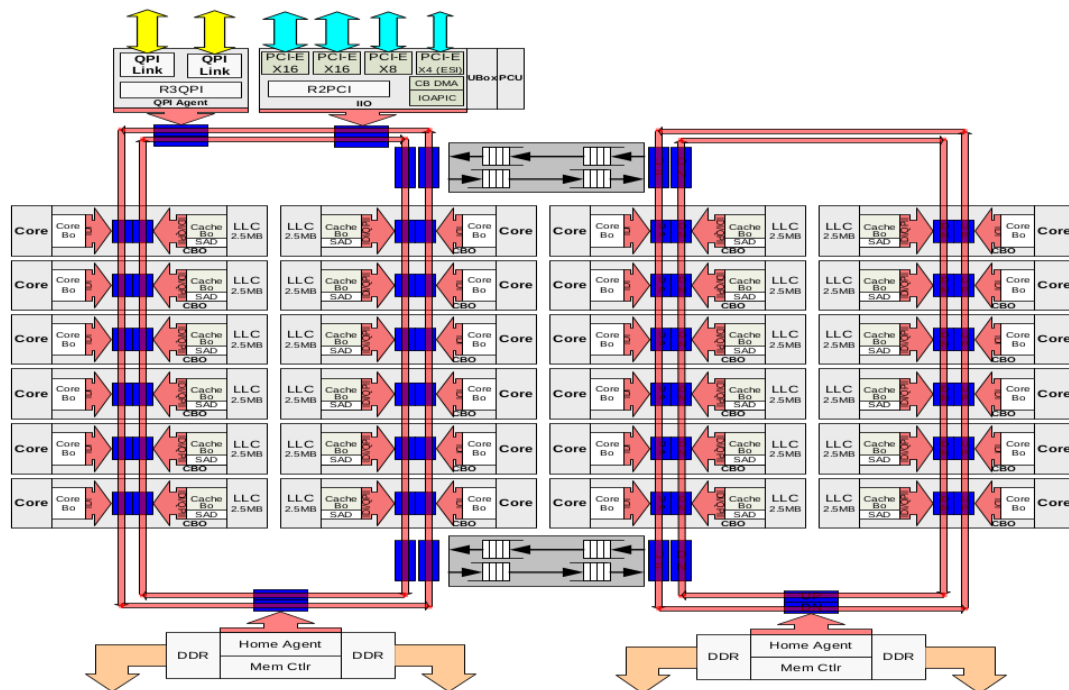
6 Memory Channels
je CPU (bisher 4)

max. 28 Cores
(bisher 24)

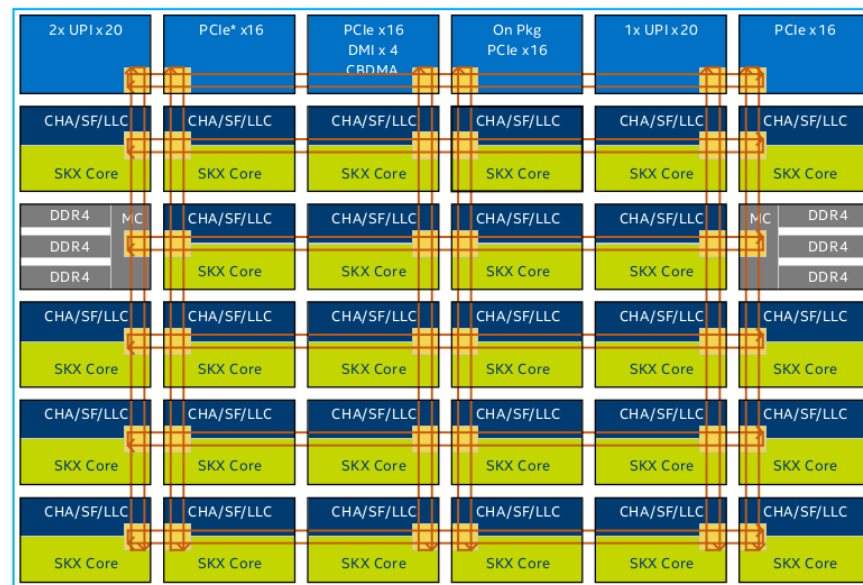
BMC: Baseboard Management Controller	PCH: Intel® Platform Controller Hub	IE: Innovation Engine
Intel® OPA: Intel® Omni-Path Architecture	Intel QAT: Intel® QuickAssist Technology	ME: Manageability Engine
NIC: Network Interface Controller	VMD: Volume Management Device	NTB: Non-Transparent Bridge

Mesh Architektur

Broadwell EX 24-core die



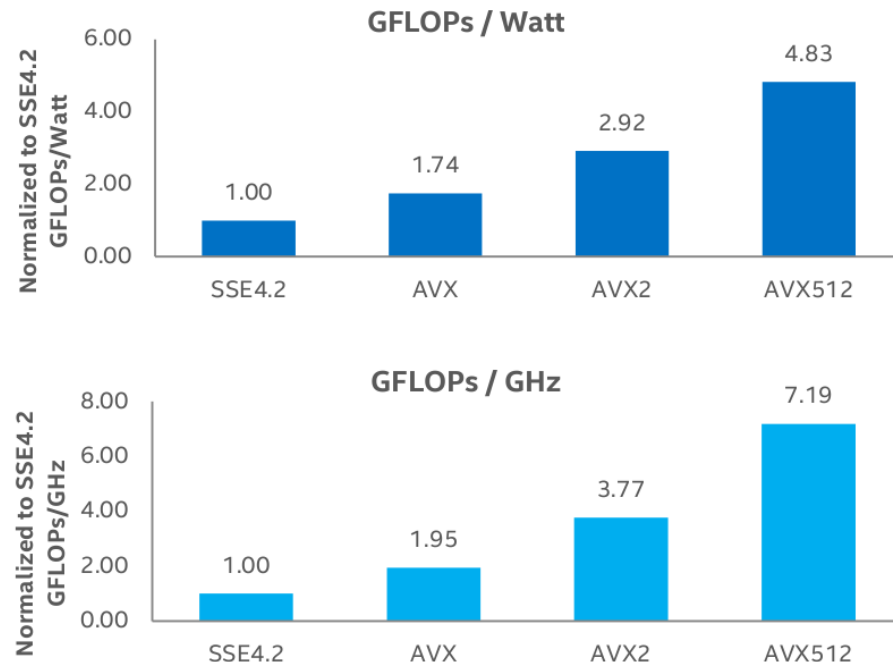
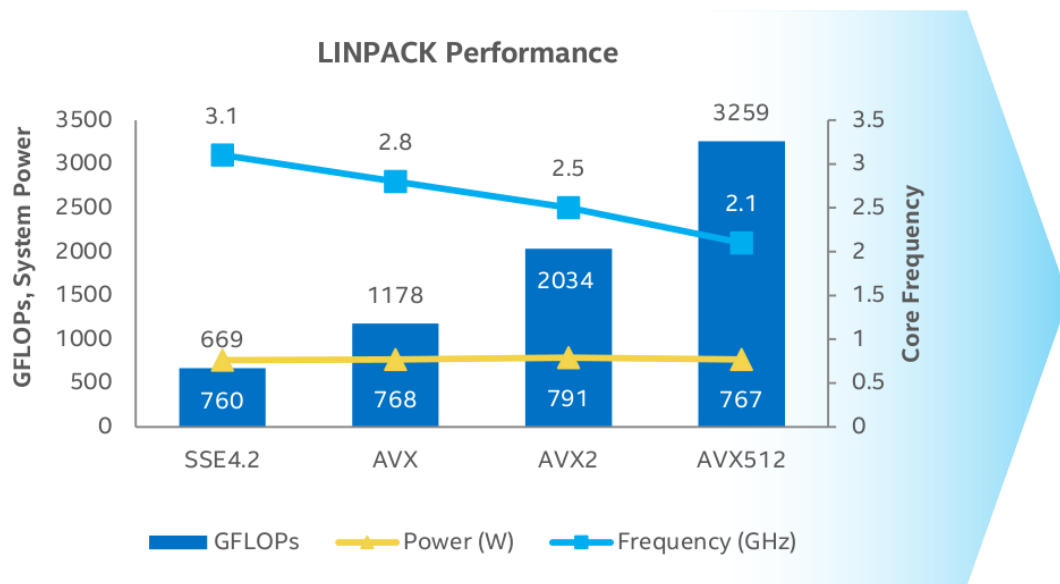
Skylake-SP 28-core die



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

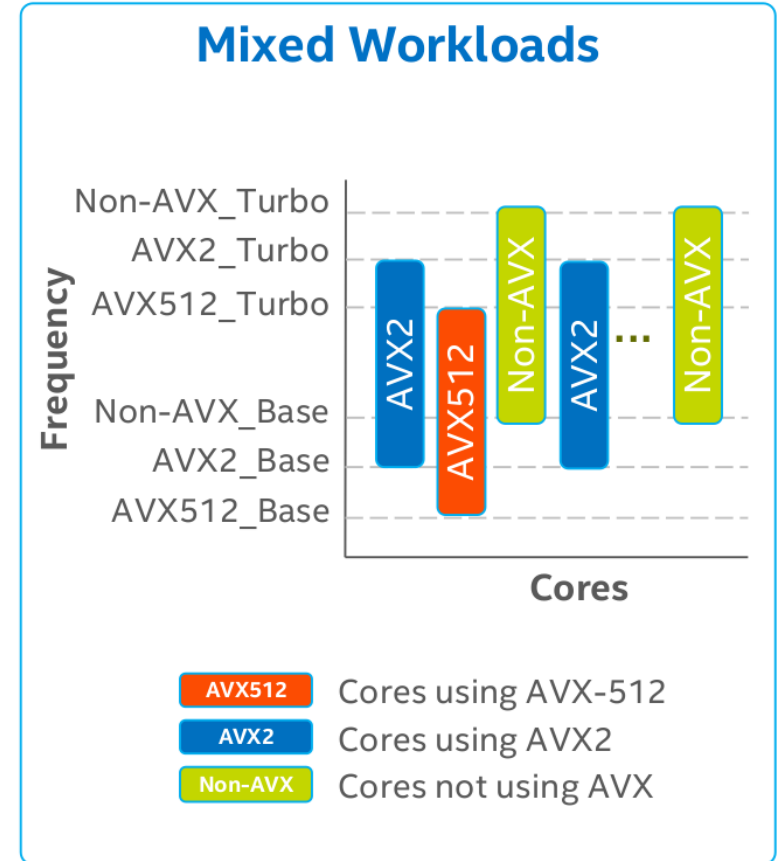
AVX-512

- LINPACK Benchmark** zur Messung der **Gleitkomma**performance:
- Misst wie schnell ein Computer ein System von **linearen Gleichungen** löst
 - Wird für **Top500** Supercomputer-Liste verwendet
 - Schwerpunkt auf CPU, nicht auf Memory oder I/O Bandbreite
 - Wird häufig auf CPU optimiert – echter Code muss vom Programmierer ebenso angepasst werden, damit Performancevorteil entsteht

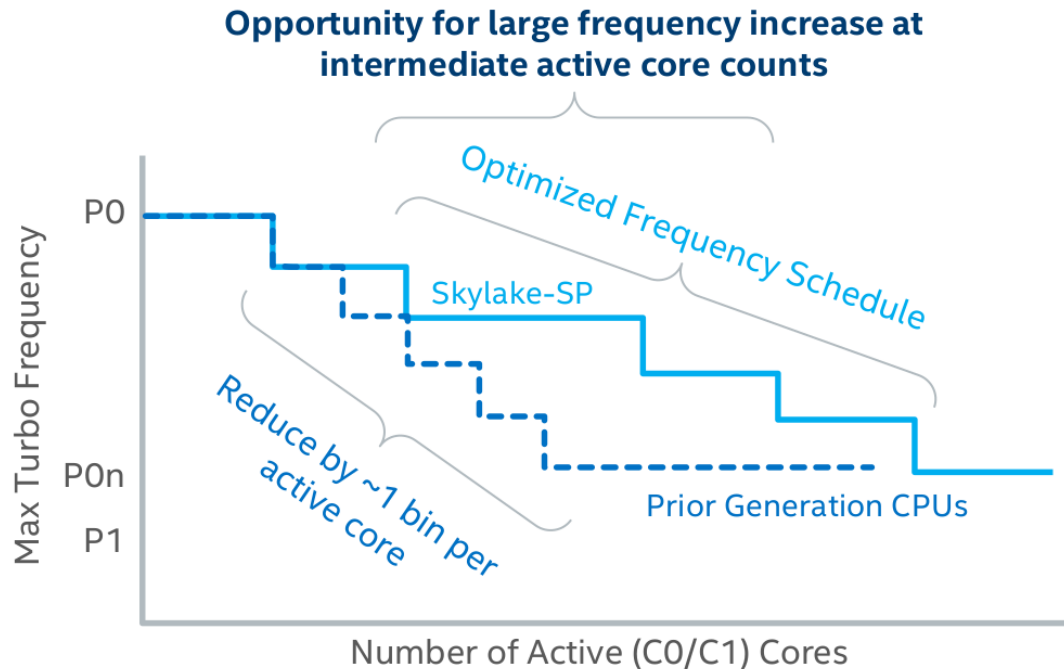


AVX-512

- FMA (Fused Multiply Add) Einheiten für AVX-512
 - 1 FMA (Bronze, Silber, Gold 5100)
 - 2 FMA (Gold 6100, Platinum)
- AVX-512 Performancevorteil abhängig von Anwendung, vor allem Forschung, Finanz, KI, Deep Learning



Optimierte Turbo Profile → höhere Taktfrequenz

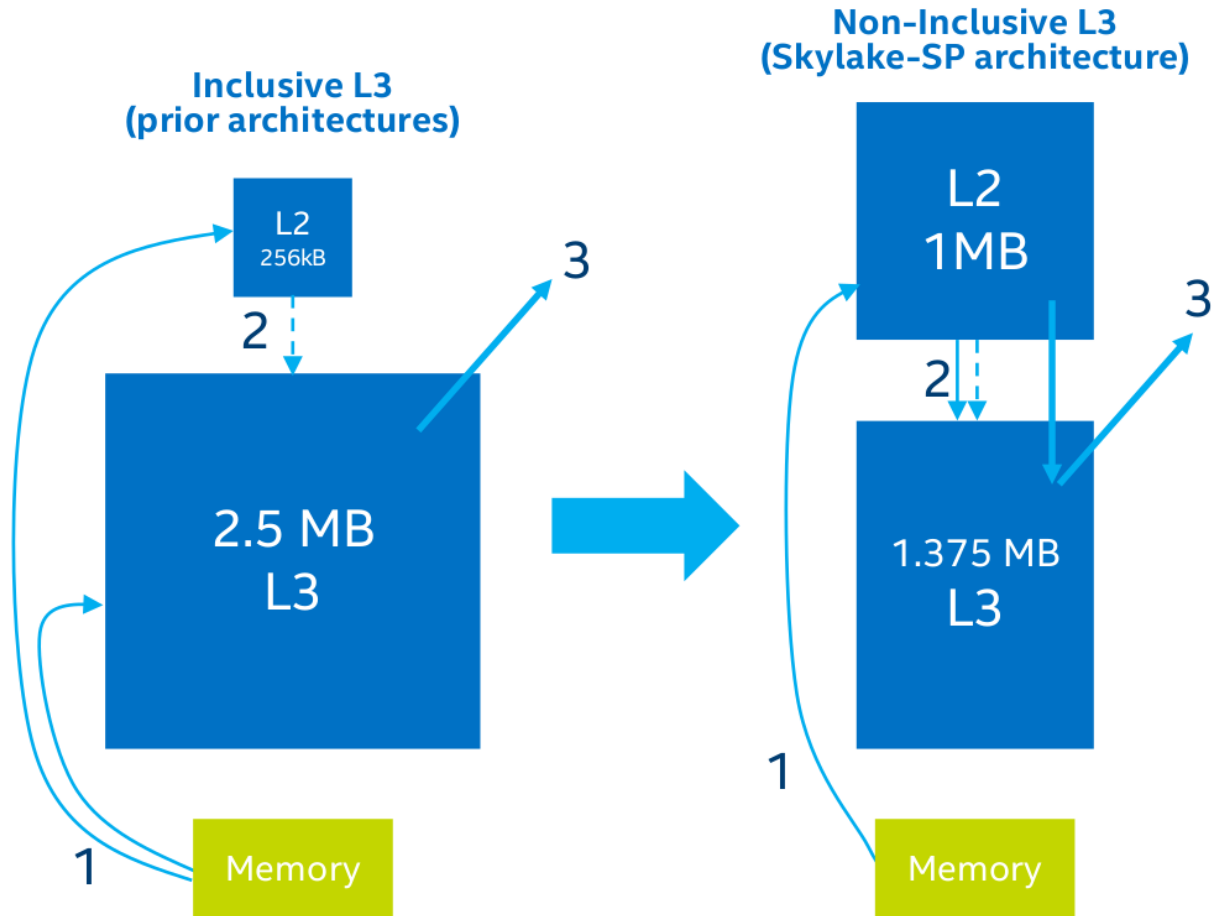


- Bisherige CPUs (Prior Generation) schalten ca. 1 Bin pro aktiven Core herunter
- Skylake-SP CPUs schalten optimiert langsamer herunter
- → Möglichkeit höherer Frequenzen bei mittlerer Anzahl aktiver Cores

Größere L2 Caches – optimiert fürs RZ



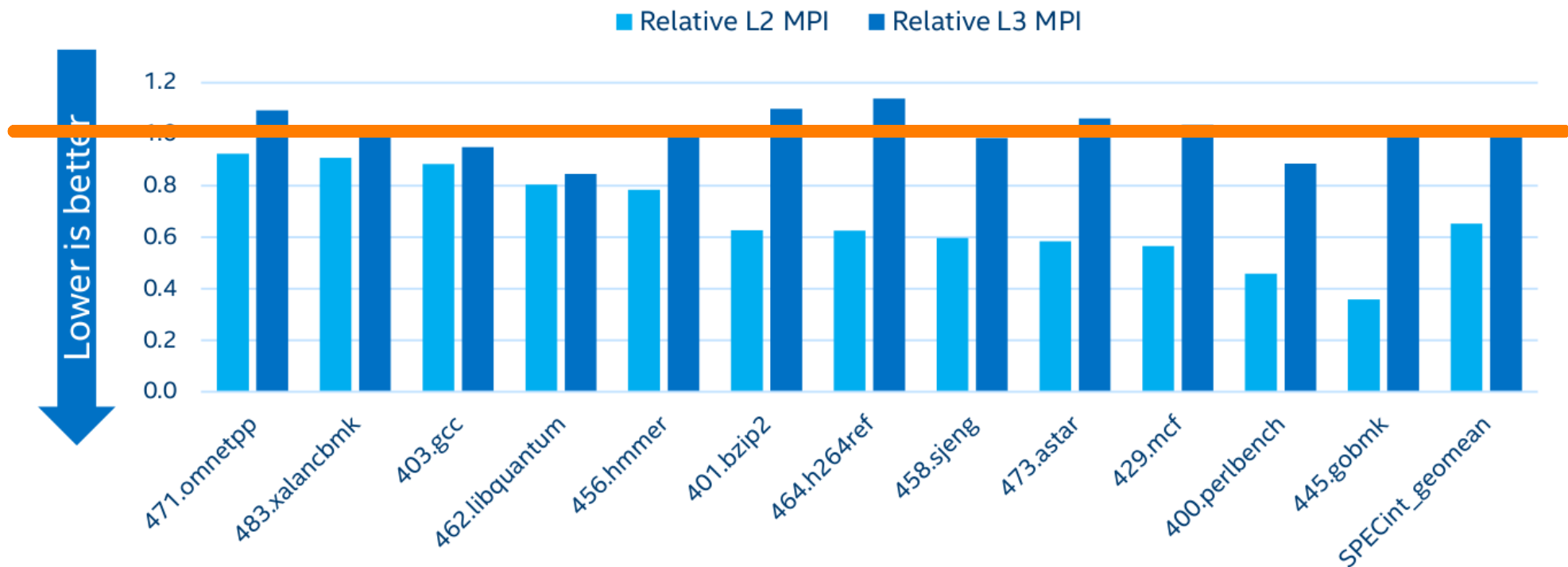
- On-Chip Balance von „shared-distributed“ → „private-local“
- Shared L3 von „inclusive“ → „non-inclusive“



- 1) Memory Read: **direkt in L2** (nicht mehr L2+L3)
- 2) Wenn Eintrag in L2 freigemacht werden muss, wird dieser in L3 übertragen (sowohl veränderte als auch unveränderte Einträge)
- 3) Daten, die von mehreren Cores genutzt werden, kommen in L3

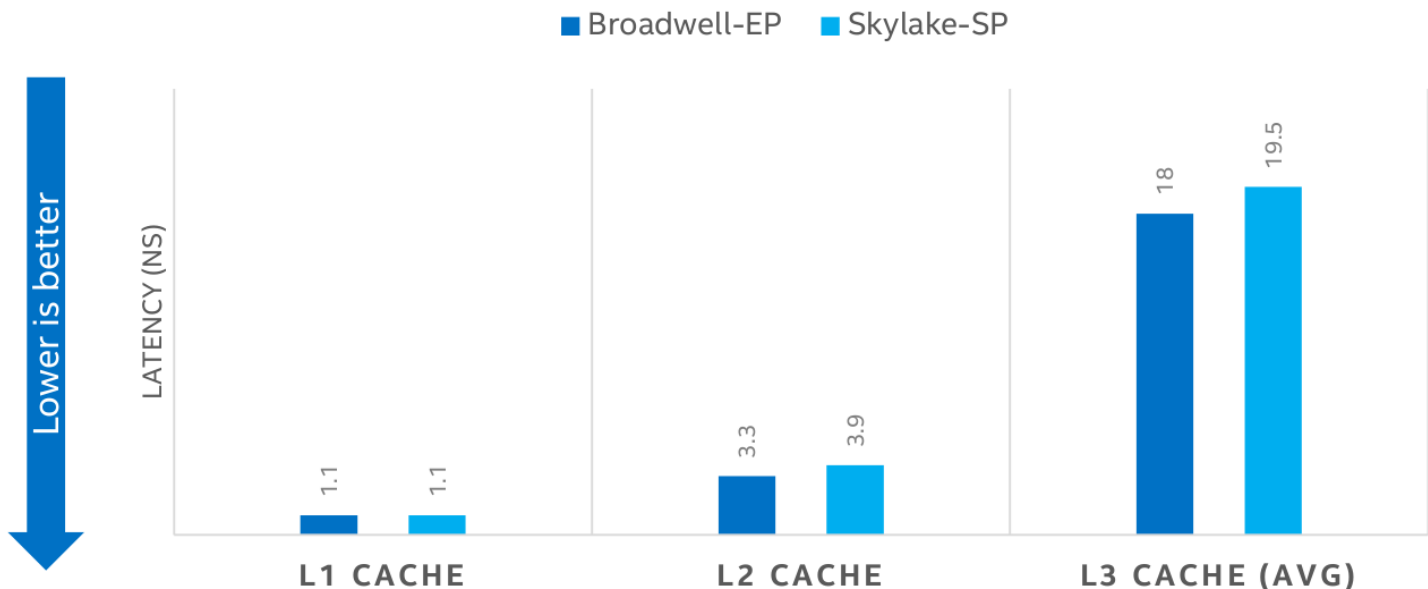
Cache Performance

Relative Change in L2 and L3 Misses Per Instruction for SPECint*_rate
2006 from Broadwell-EP to Skylake-SP



Cache Performance

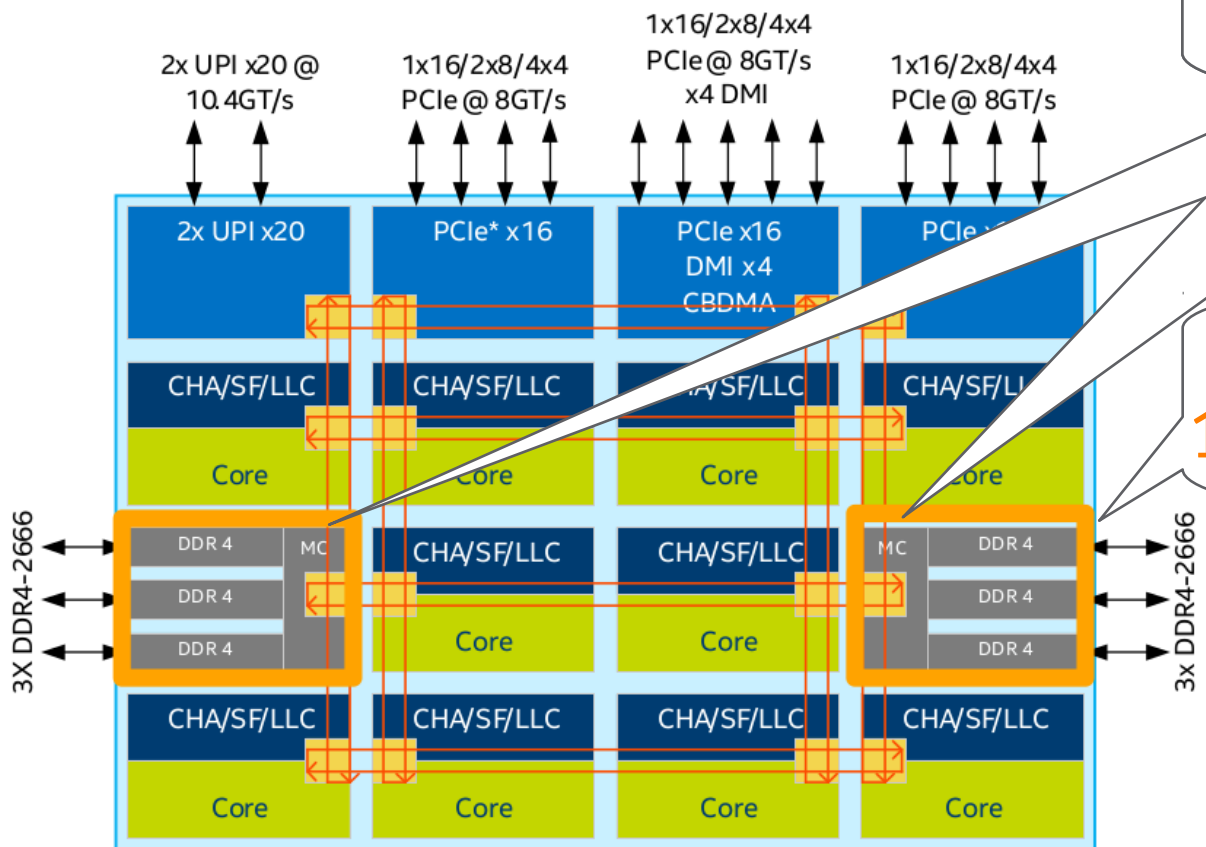
CPU CACHE LATENCY



- Skylake-SP Latenz etwas erhöht – für **4x so viel L2**
- Skylake-SP gute L3 Cache Latenzen bei **mehr Cores (28 vs. 24)**

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.

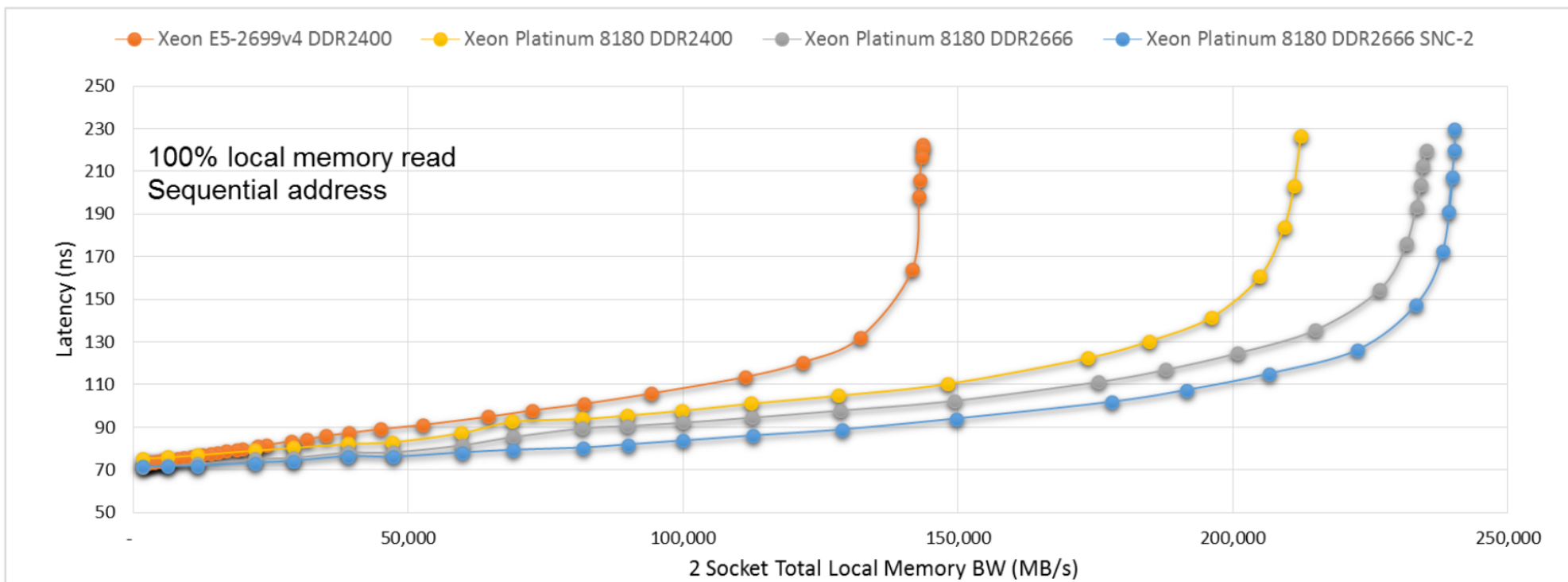
Memory Subsystem



2 Memory Controller
mit je 3 Channels

→ 6 Channels pro CPU,
12 Channels bei Dual-CPU

Memory Performance (Bandbreite/Latenz)



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>

Regeln für optimale Speicher-Performance

- 1) **Kanäle:** Alle bestückten Memory Channels sollten die selbe Gesamtkapazität (GB) und die selbe Anzahl an Ranks haben

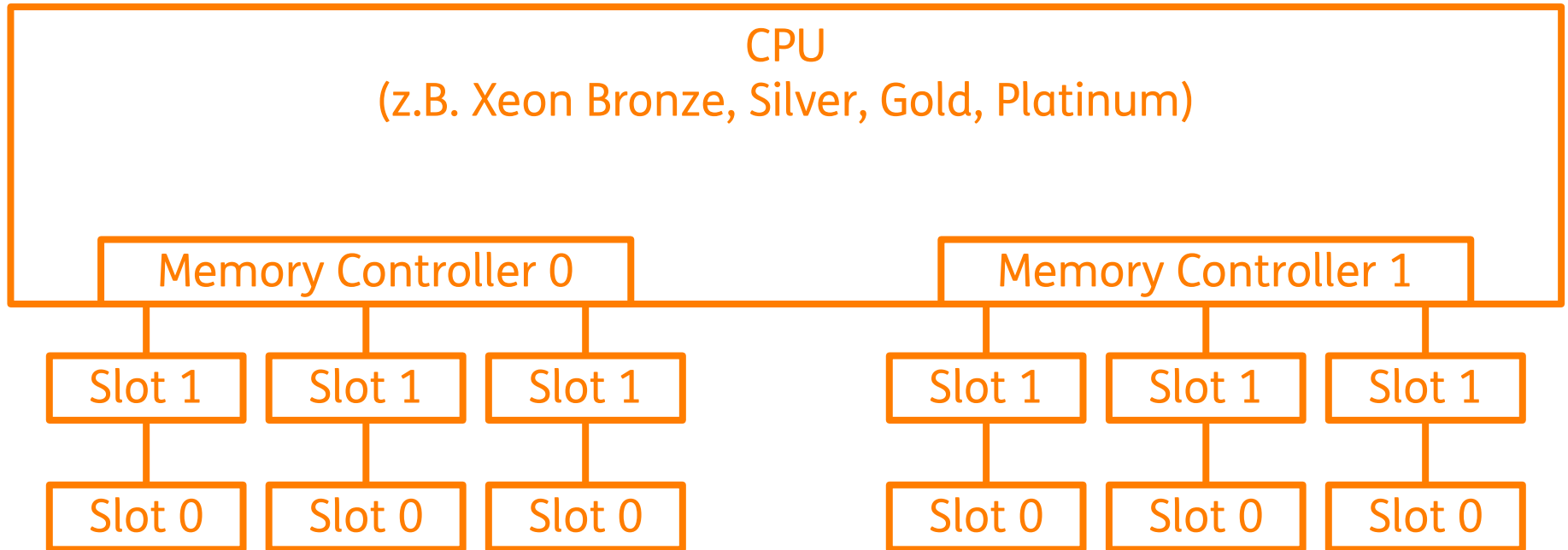
Regeln für optimale Speicher-Performance

- 1) **Kanäle:** Alle bestückten Memory Channels sollten die selbe Gesamtkapazität (GB) und die selbe Anzahl an Ranks haben
- 2) **Controller:** Alle (beide) Memory Controller einer CPU sollten die selbe Konfiguration von RAM Riegeln (DIMMs) haben

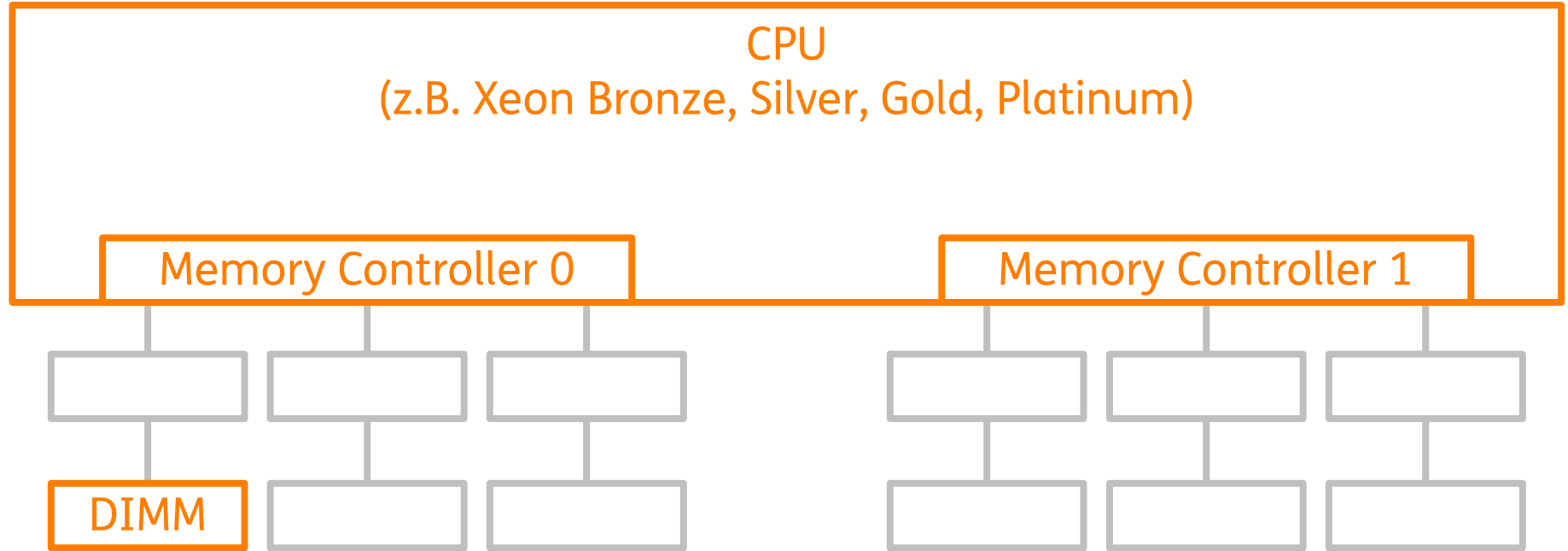
Regeln für optimale Speicher-Performance

- 1) **Kanäle:** Alle bestückten Memory Channels sollten die selbe Gesamtkapazität (GB) und die selbe Anzahl an Ranks haben
- 2) **Controller:** Alle (beide) Memory Controller einer CPU sollten die selbe Konfiguration von RAM Riegeln (DIMMs) haben
- 3) **CPUs:** Alle CPUs eines Servers (z.B. beide CPUs eines Dual-CPU Servers) sollten die selbe Konfiguration RAM Riegeln (DIMMs) haben

Optimale Speicher-Performance – Beispiele



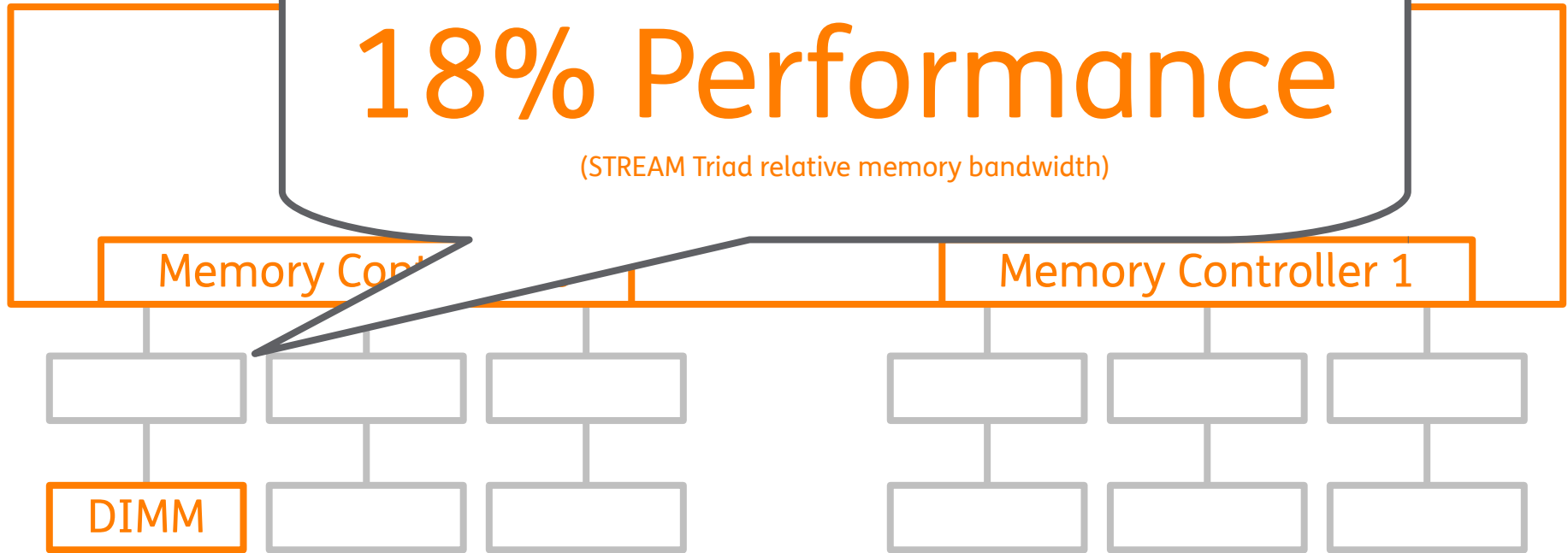
Optimale Speicher-Performance – Beispiele



Optimale Speicher-Performance – Beispiele

18% Performance

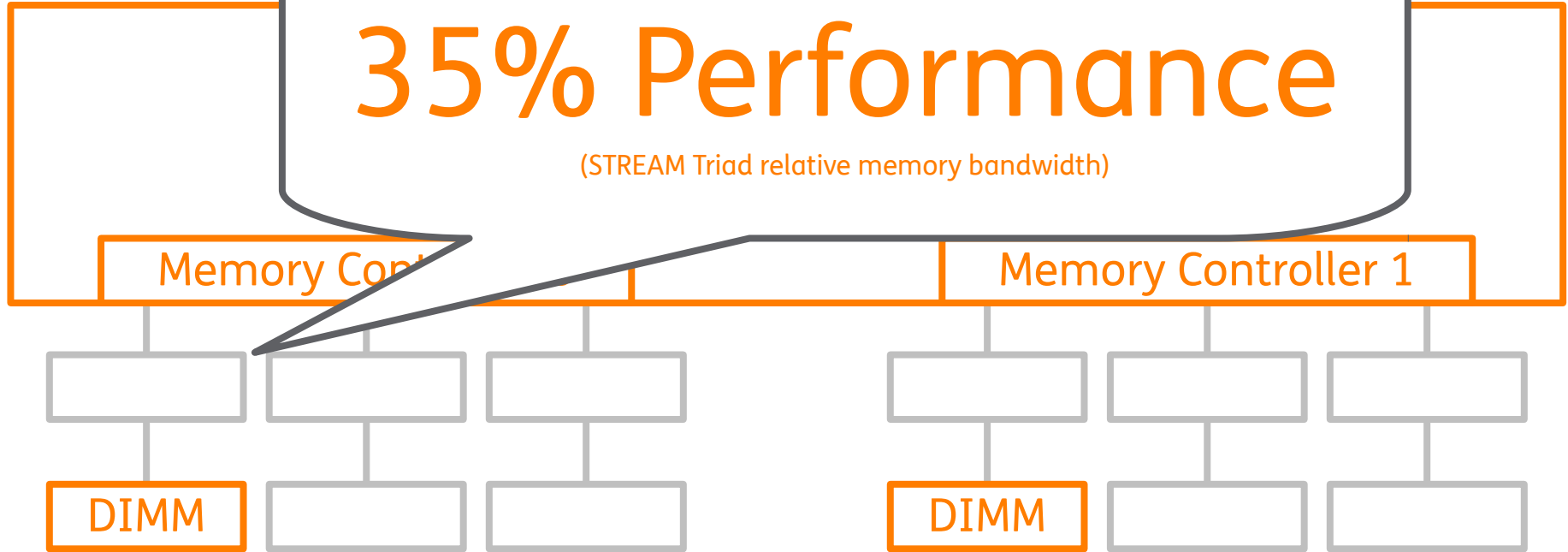
(STREAM Triad relative memory bandwidth)



Optimale Speicher-Performance – Beispiele

35% Performance

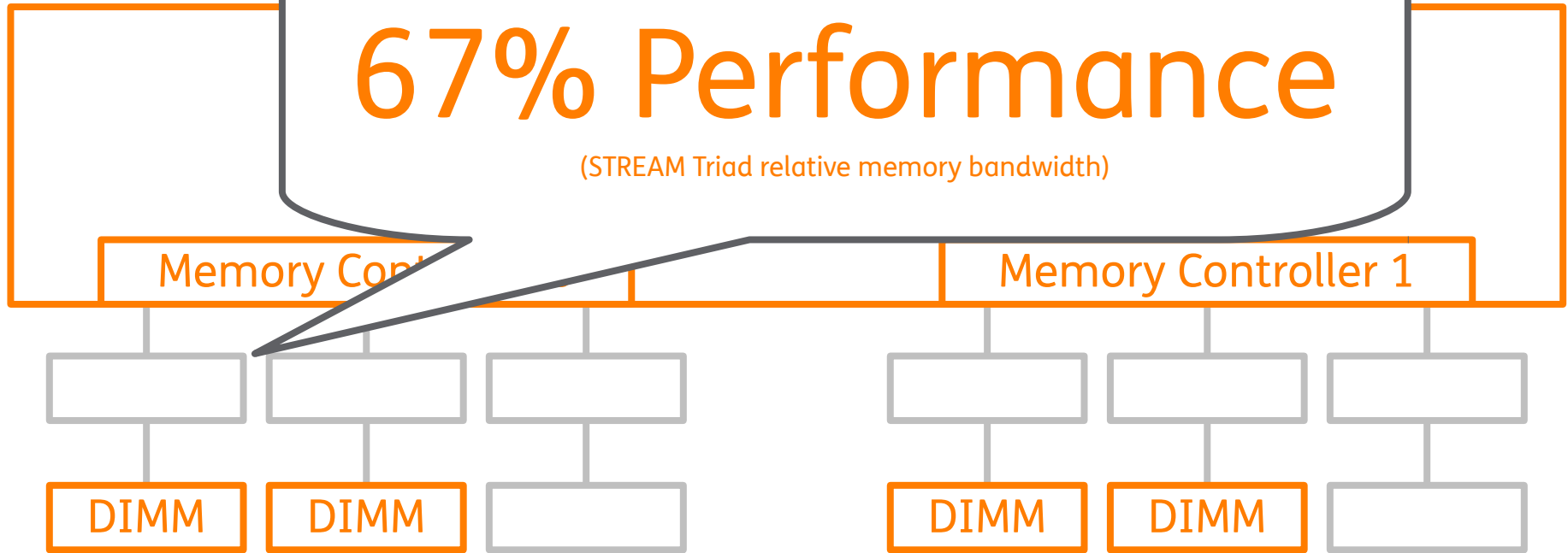
(STREAM Triad relative memory bandwidth)



Optimale Speicher-Performance – Beispiele

67% Performance

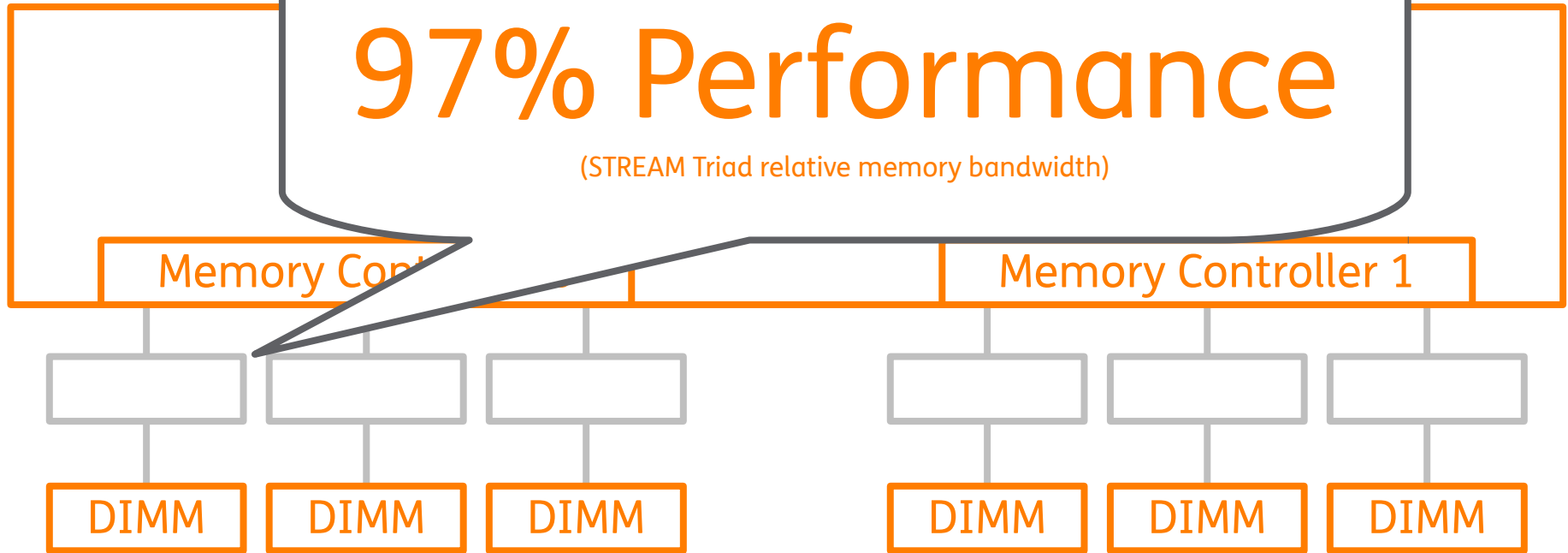
(STREAM Triad relative memory bandwidth)



Optimale Speicher-Performance – Beispiele

97% Performance

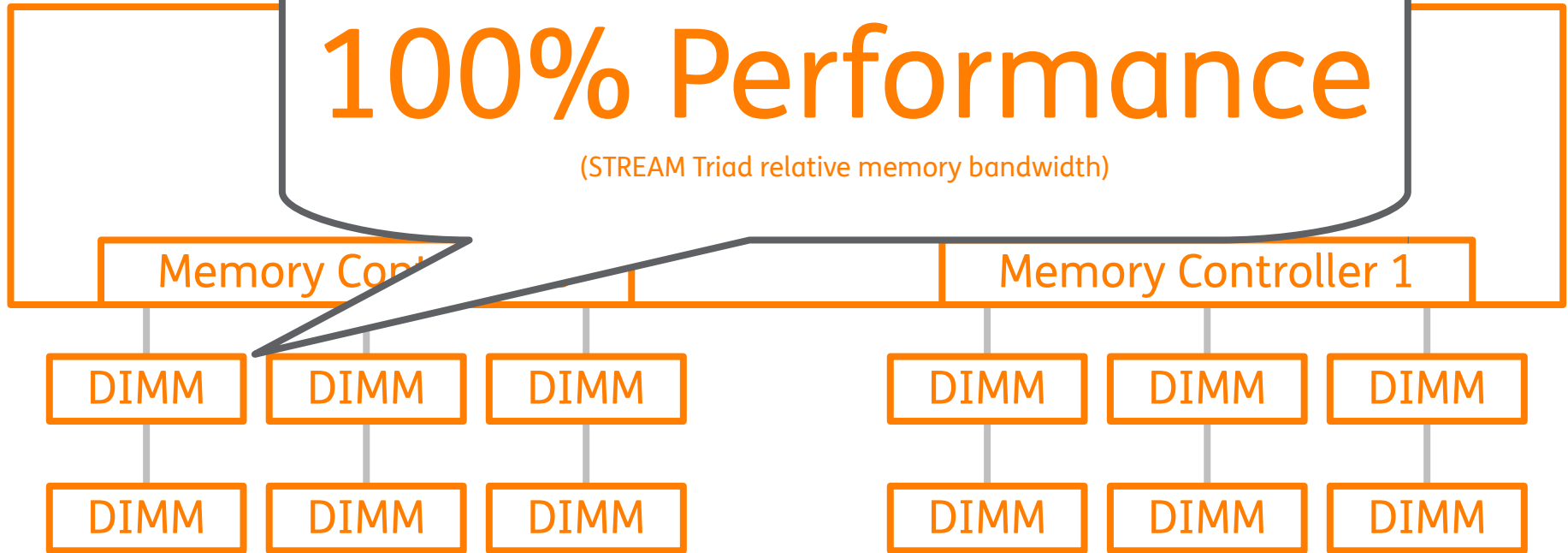
(STREAM Triad relative memory bandwidth)



Optimale Speicher-Performance – Beispiele

100% Performance

(STREAM Triad relative memory bandwidth)



RAS (Reliability, Availability, Serviceability)

- Neue und erweiterte RAS Funktionen:
- Standard RAS (Bronze/Silver)
- Advanced RAS (Gold/Platinum)

RAS (Reliability, Availability, Serviceability)

- Adaptive Data Correction (ADC)
- Adaptive Double DRAM Device Correction (ADDDC) – Advanced
- Advanced Error Detection and Correction (AEDC)
- LMCE (4&8 Socket) – Advanced
 - ab Linux Kernel 4.2/4.3, in KVM Gästen ab Kernel 4.8 (siehe [Git Commits](#))

Intel Xeon Scalable Performance

– Das ist NEU:

- Neue Mikroarchitektur
- AVX512
- Größere L2-Caches
- Memory Subsystem – 6 Channels/CPU
- RAS Funktionen

– Die Prozessoren: Xeon Bronze, Silver, Gold und Platinum

– Skylake-SP Server bei Thomas-Krenn





CPU	3100 Serie	4100 Serie	5100 Serie	6100 Serie	8100 Serie
Cores	6-8	4-12	4-14	6-22	4-28
Max. Taktfrequenz	1,7 GHz (8C)	2,2 GHz (10C)	3,6 GHz (4C)	3,4 GHz (6C)	3,6 GHz (4C)
AVX-512	1 FMA	1 FMA	1 FMA	2 FMAs	2 FMAs
UPI Links	2	2	2	3 ¹	3 ¹
UPI Geschwindigkeit	9,6 GT/s	9,6 GT/s	10,4 GT/s	10,4 GT/s	10,4 GT/s
DDR4 Support MT/s	2133	2400 ²	2400 ²	2666 ²	2666 ²
Hyperthreading		✓	✓	✓	✓
Turbo Boost 2.0		✓	✓	✓	✓
Turbo Max. Freq.		3,0 GHz (alle)	3,7 GHz (4C)	4,2 GHz (12C)	3,8 GHz (28C)
RAS	Standard	Standard	Advanced	Advanced	Advanced

1) abhängig von Mainboard 2) abhängig von den verwendeten RAM-Riegeln



CPU	3100 Serie	4100 Serie	5100 Serie	6100 Serie	8100 Serie
Cores	6-8	4-12	4-14	6-22	4-28
Max. Taktfrequenz	1,7 GHz (8C)	2,2 GHz (10C)	3,6 GHz (4C)	3,4 GHz (6C)	3,6 GHz (4C)
AVX-512	1 FMA	1 FMA	1 FMA	2 FMAs	2 FMAs
UPI Links	2	2	2	3 ¹	3 ¹
UPI Geschwindigkeit	9,6 GT/s	9,6 GT/s	10,4 GT/s	10,4 GT/s	10,4 GT/s
DDR4 Support MT/s	2133	2400 ²	2400 ²	2666 ²	2666 ²
Hyperthreading		✓	✓	✓	✓
Turbo Boost 2.0		✓	✓	✓	✓
Turbo Max. Freq.		3,0 GHz (alle)	3,7 GHz (4C)	4,2 GHz (12C)	3,8 GHz (28C)
RAS	Standard	Standard	Advanced	Advanced	Advanced

1) abhängig von Mainboard 2) abhängig von den verwendeten RAM-Riegeln



CPU	3100 Serie	4100 Serie	5100 Serie	6100 Serie	8100 Serie
Cores	6-8	4-12	4-14	6-22	4-28
Max. Taktfrequenz	1,7 GHz (8C)	2,2 GHz (10C)	3,6 GHz (4C)	3,4 GHz (6C)	3,6 GHz (4C)
AVX-512	1 FMA	1 FMA	1 FMA	2 FMAs	2 FMAs
UPI Links	2	2	2	3 ¹	3 ¹
UPI Geschwindigkeit	9,6 GT/s	9,6 GT/s	10,4 GT/s	10,4 GT/s	10,4 GT/s
DDR4 Support MT/s	2133	2400 ²	2400 ²	2666 ²	2666 ²
Hyperthreading		✓	✓	✓	✓
Turbo Boost 2.0		✓	✓	✓	✓
Turbo Max. Freq.		3,0 GHz (alle)	3,7 GHz (4C)	4,2 GHz (12C)	3,8 GHz (28C)
RAS	Standard	Standard	Advanced	Advanced	Advanced

1) abhängig von Mainboard 2) abhängig von den verwendeten RAM-Riegeln



CPU	3100 Serie	4100 Serie	5100 Serie	6100 Serie	8100 Serie
Cores	6-8	4-12	4-14	6-22	4-28
Max. Taktfrequenz	1,7 GHz (8C)	2,2 GHz (10C)	3,6 GHz (4C)	3,4 GHz (6C)	3,6 GHz (4C)
AVX-512	1 FMA	1 FMA	1 FMA	2 FMAs	2 FMAs
UPI Links	2	2	2	3 ¹	3 ¹
UPI Geschwindigkeit	9,6 GT/s	9,6 GT/s	10,4 GT/s	10,4 GT/s	10,4 GT/s
DDR4 Support MT/s	2133	2400 ²	2400 ²	2666 ²	2666 ²
Hyperthreading		✓	✓	✓	✓
Turbo Boost 2.0		✓	✓	✓	✓
Turbo Max. Freq.		3,0 GHz (alle)	3,7 GHz (4C)	4,2 GHz (12C)	3,8 GHz (28C)
RAS	Standard	Standard	Advanced	Advanced	Advanced

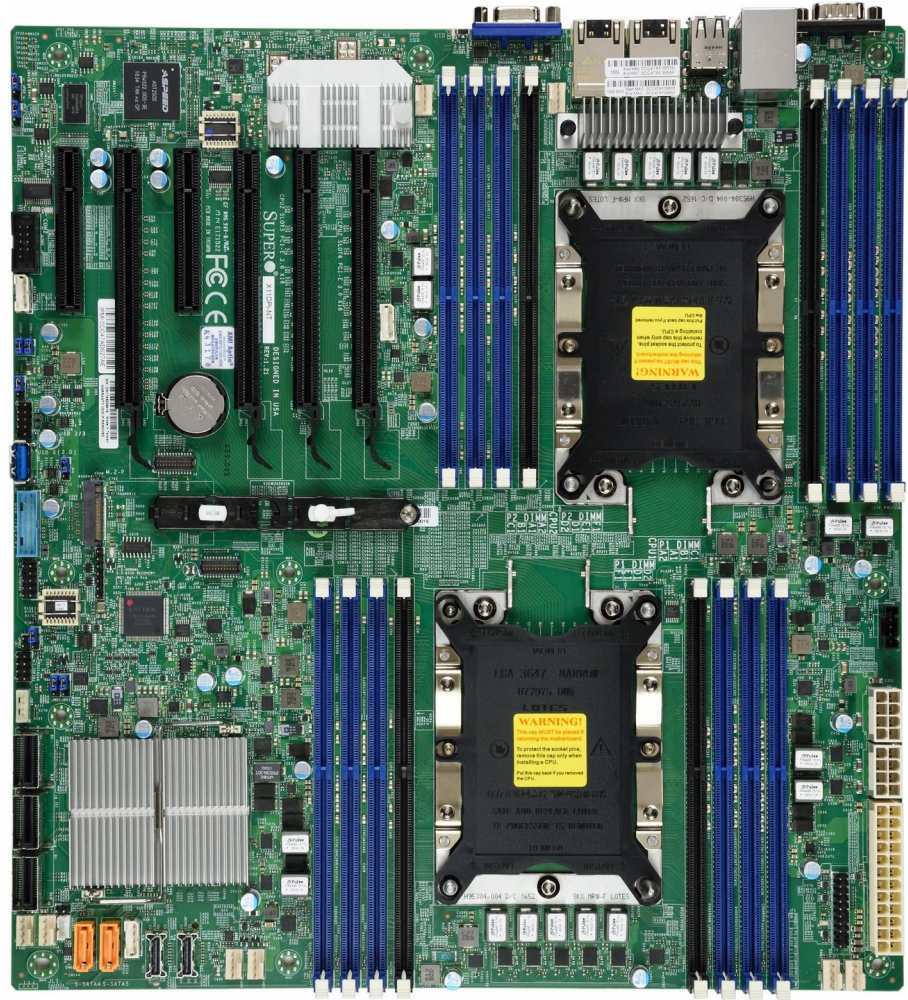
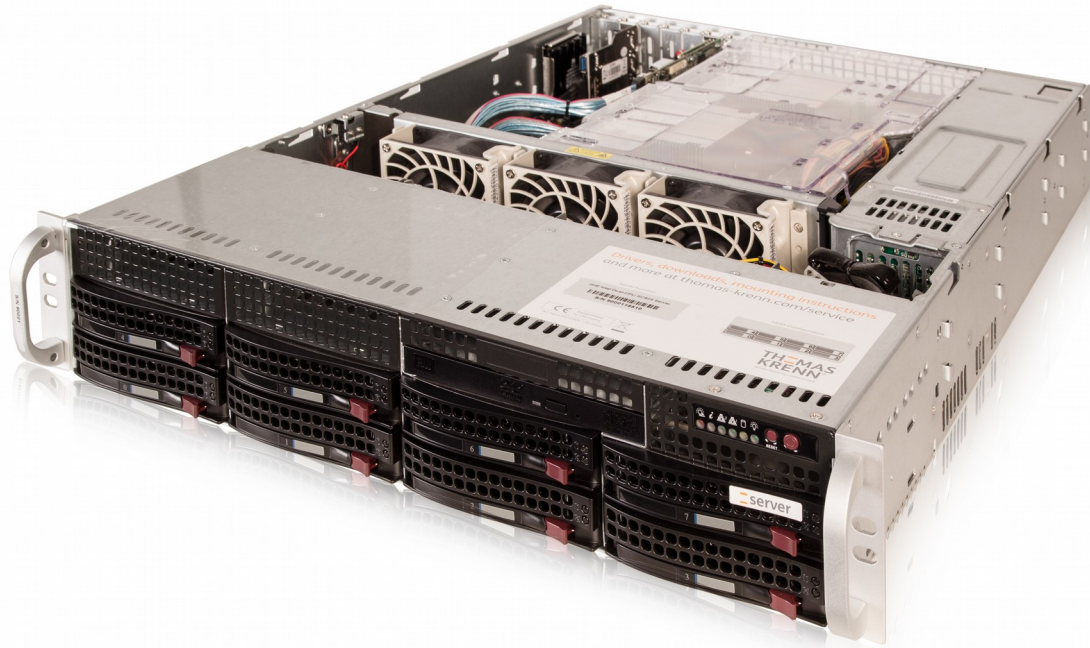
1) abhängig von Mainboard 2) abhängig von den verwendeten RAM-Riegeln



CPU	3100 Serie	4100 Serie	5100 Serie	6100 Serie	8100 Serie
Cores	6-8	4-12	4-14	6-22	4-28
Max. Taktfrequenz	1,7 GHz (8C)	2,2 GHz (10C)	3,6 GHz (4C)	3,4 GHz (6C)	3,6 GHz (4C)
AVX-512	1 FMA	1 FMA	1 FMA	2 FMAs	2 FMAs
UPI Links	2	2	2	3 ¹	3 ¹
UPI Geschwindigkeit	9,6 GT/s	9,6 GT/s	10,4 GT/s	10,4 GT/s	10,4 GT/s
DDR4 Support MT/s	2133	2400 ²	2400 ²	2666 ²	2666 ²
Hyperthreading		✓	✓	✓	✓
Turbo Boost 2.0		✓	✓	✓	✓
Turbo Max. Freq.		3,0 GHz (alle)	3,7 GHz (4C)	4,2 GHz (12C)	3,8 GHz (28C)
RAS	Standard	Standard	Advanced	Advanced	Advanced

Intel Xeon Scalable Performance

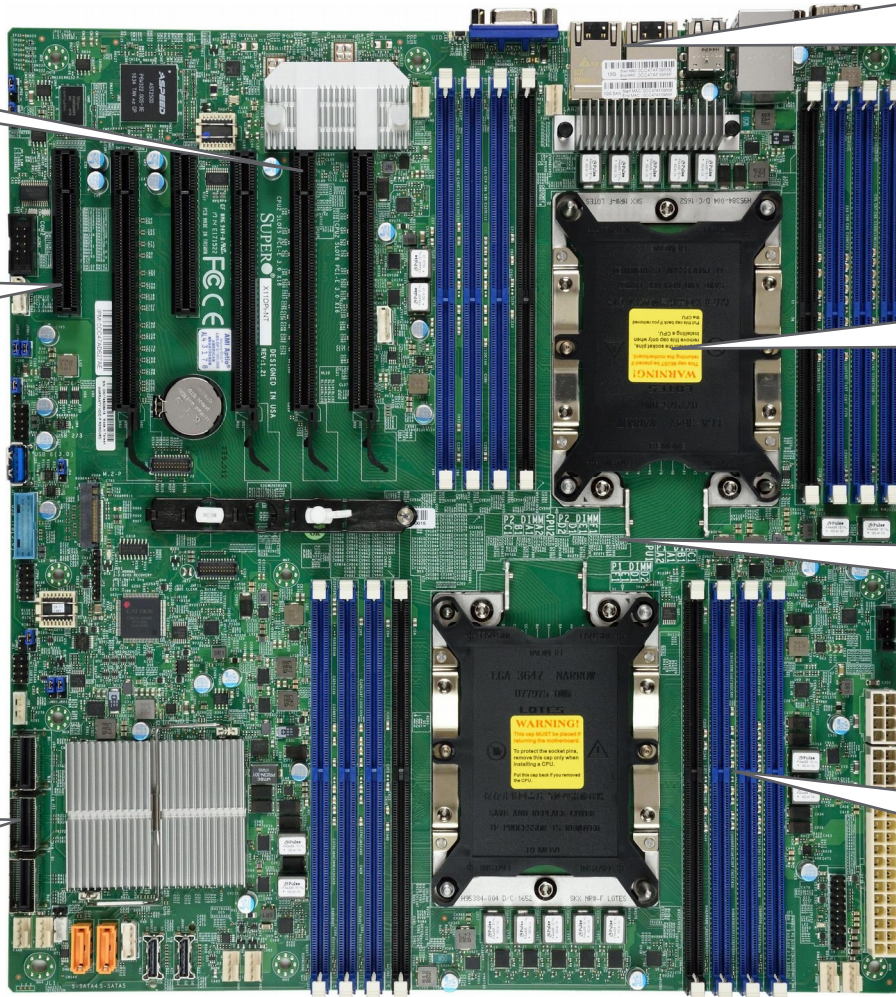
- Das ist NEU:
 - Neue Mikroarchitektur
 - AVX512
 - Größere L2-Caches
 - Memory Subsystem – 6 Channels/CPU
 - RAS Funktionen
- Die Prozessoren: Xeon Bronze, Silver, Gold und Platinum
- Skylake-SP Server bei Thomas-Krenn



4 x PCIe 3.0 x16

2 x PCIe 3.0 x8

14 (!) x SATA3



2 x LAN
(1 oder 10Gbit)

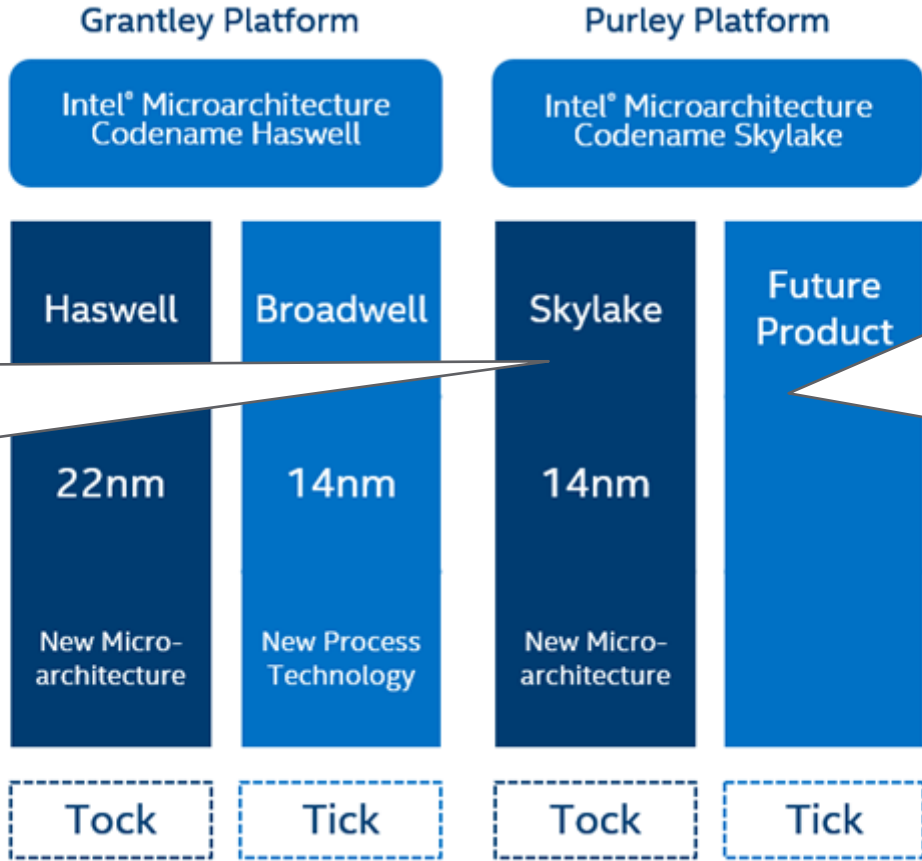
2 CPU Sockets

2 UPI Links

16 DIMM Slots

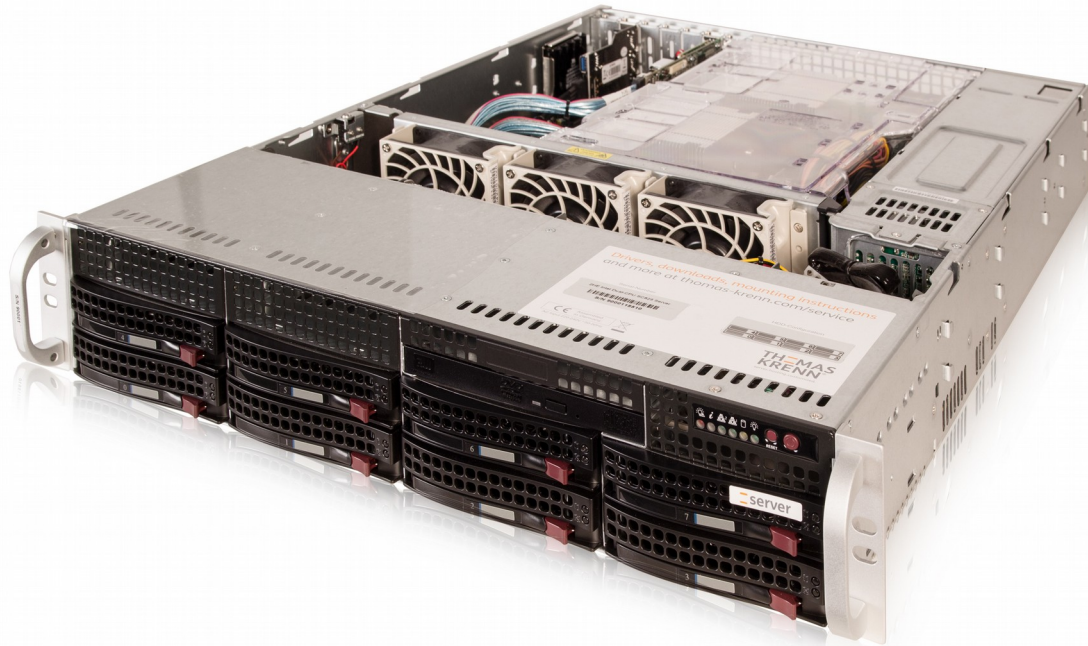
Ausblick

Xeon Scalable Platform
Bronze 3100,
Silber 4100,
Gold 5100/6100,
Platinum 8100



Xeon Scalable family refresh "Cascade Lake"¹
(Ende?) 2018,
NVDIMMs,
pmem.io

1) Quelle: Intel Pressemitteilung "Intel Shows 1.59x Performance Improvement in Upcoming Intel Xeon Processor Scalable Family" (16.05.2017)



**Hohe Memory
Performance: +60%**



**Level 2 Cache: +300%
(4x so viel wie bisher)**



**Optimierte
Turbo Boost Profile**

**THOMAS
KRENN®**



Intel Xeon Scalable Performance Systeme:

verfügbar in Kürze



**THOMAS
KRENN®**