Intel's High-End
Server Processors and Platforms

Dezső Sima

January 2018

(Ver. 2.0)

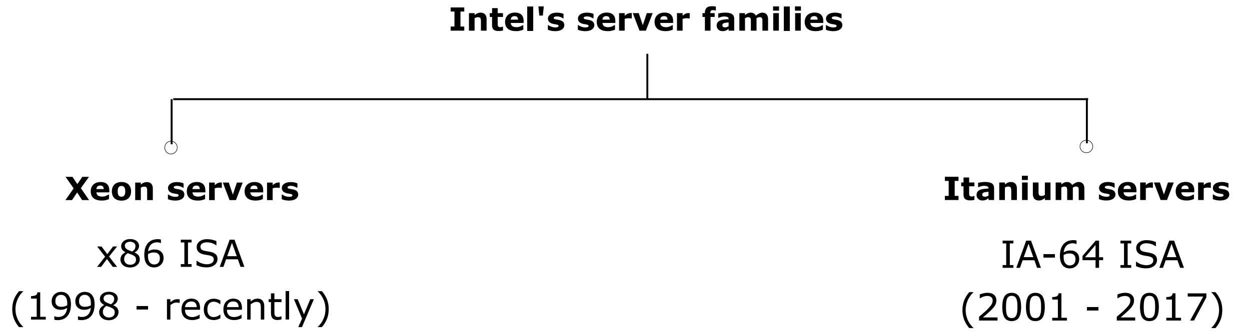
© Sima Dezső, 2018

Intel’s high-end 4S/8S server processors and platforms

- 1. Introduction to Intel’ s high-end multicore 4S/8S server platforms
- 2. Evolution of Intel’s high-end multicore 4S/8S server platforms
- 3. Example 1: The Brickland platform
- 4. Example 2: The Purley platform
- 5. References

Remark

The material presented in these slides is restricted to Intel’s high-end multicore x86 server (called Xeon) processors and platforms, accordingly, previous and IA64 (Itanium) based processors and systems are not covered.



Intel's first multicore server processors were based on the 3. core of the Pentium 4 family (called Prescott), there were introduced about 2005.

1. Introduction to Intel's high-end multicore 4S/8S server platforms

- 1.1 The worldwide server market
- 1.2 The platform concept
- 1.3 Server platforms classified according to their scalability
- 1.4 Multiprocessor server platforms classified according to their memory architecture
- 1.5 Platforms classified according to the number of chips constituting the chipset
- 1.6 Naming scheme of Intel's server processors
- 1.7 Overview of Intel's high-end 4S/8S multicore server platforms

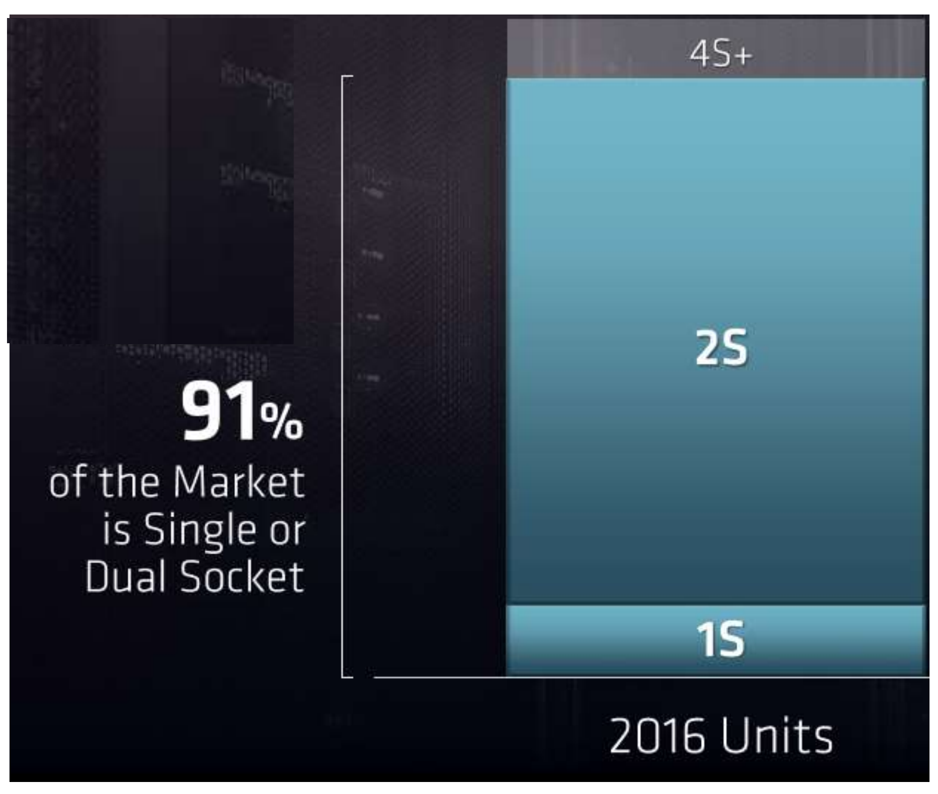
1.1 The worldwide server market

1. Introduction to Intel’s high-end multicore 4S/8S server platforms

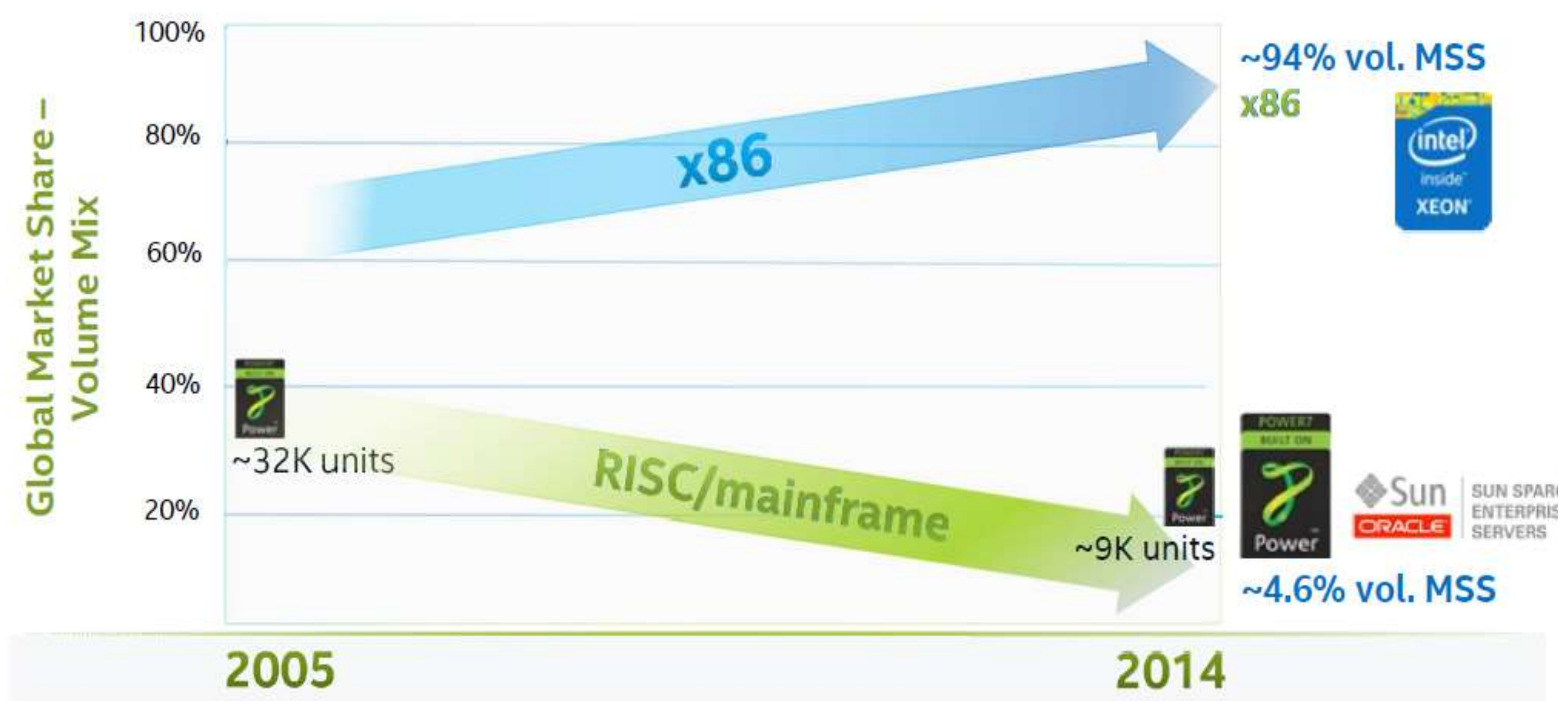
Figure: Recent datacenter [175]

1.1 The worldwide server market (1)

Recent market share of 1S and 2S servers [170]

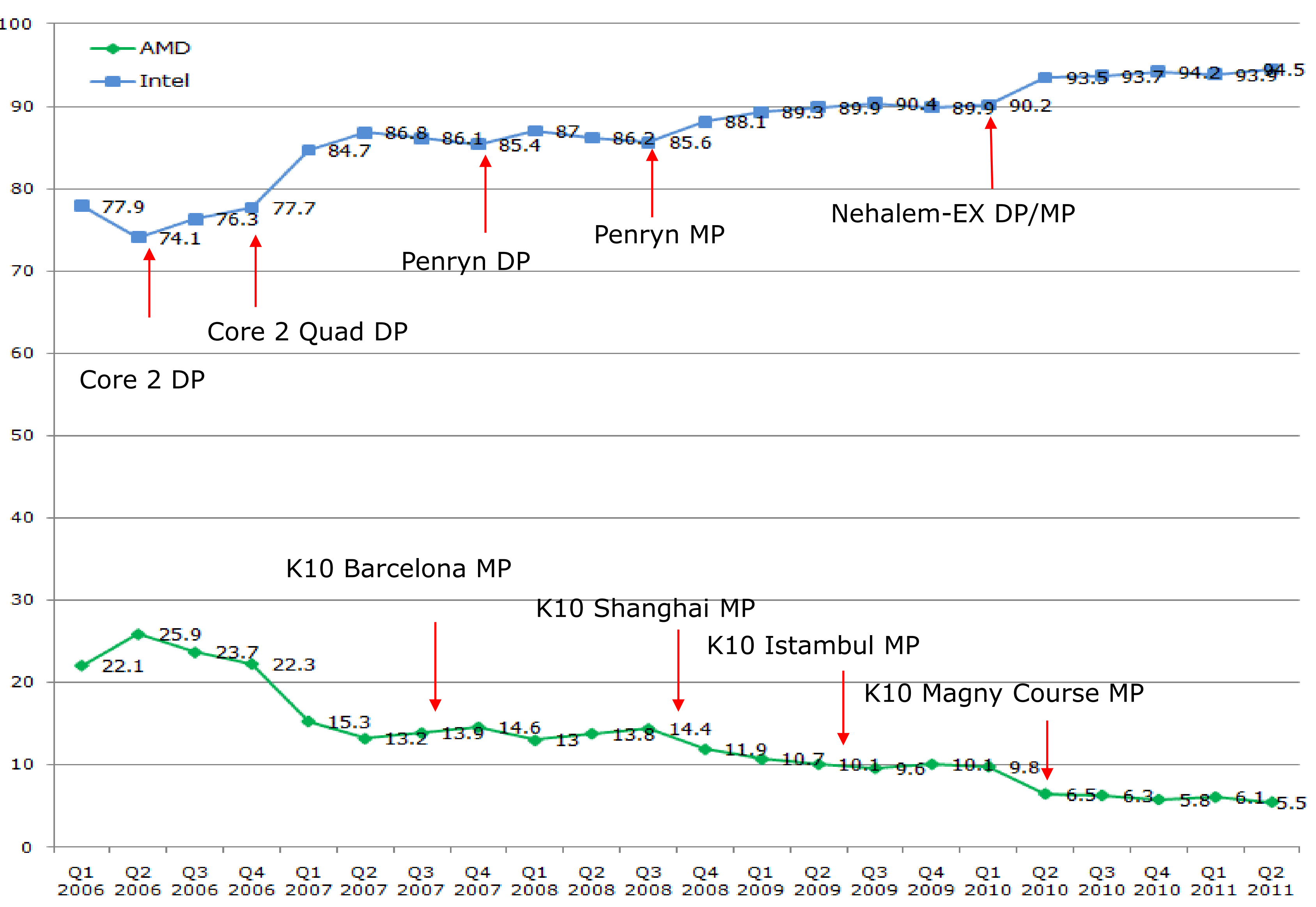


The worldwide 4S (4-socket) server market [52]



MSS: Market Segment Share

Intel's vs AMD's x86 server market share (Based on data by IDC, Mercury Research) [171]



1.1 The worldwide server market (4)

Remark

- In 06/2017 AMD launched their Epyc server processor line supporting at its introduction 2-socket configurations.

  Epyc server processors are based on the Zen microarchitecture.
- Time will show how far AMD will be able to gain market share against Intel's latest, Skylake-SP based servers.

*

Top 5 worldwide server systems vendor revenues, market shares and growth, Q1 2016, (Revenues are in Millions USD) [137]

Vendor	1Q16 Revenue	1Q16 Market Share	1Q16/1Q15 Revenue Growth
1. HP	$3,306.8	26.7%	3.5%
2. Dell	$2,267.8	18.3%	-1.8%
3. IBM	$1,139.5	9.2%	-32.9%
4. Lenovo *	$871.2	7.0%	-8.6%
5. Cisco *	$850.2	6.9%	-4.5%
ODM Direct	$863.8	7.0%	-11.0%
Others	$3,082.4	24.9%	9.0%
Total	**$12,382**	**100%**	**-3.6%**

Source: IDC's Worldwide Quarterly Server Tracker, June 2016

HPE: HP Enterprise

1.2 The platform concept

1.2 The platform concept

The notion platform is widely used in different segments of the IT industry e.g. by IC manufacturers, system providers or even by software suppliers with different interpretations.

Here we are focusing on the platform concept as used by system providers, like Intel or AMD.

The platform concept of system providers

The platform is defined by the main components of the system architecture, being typically

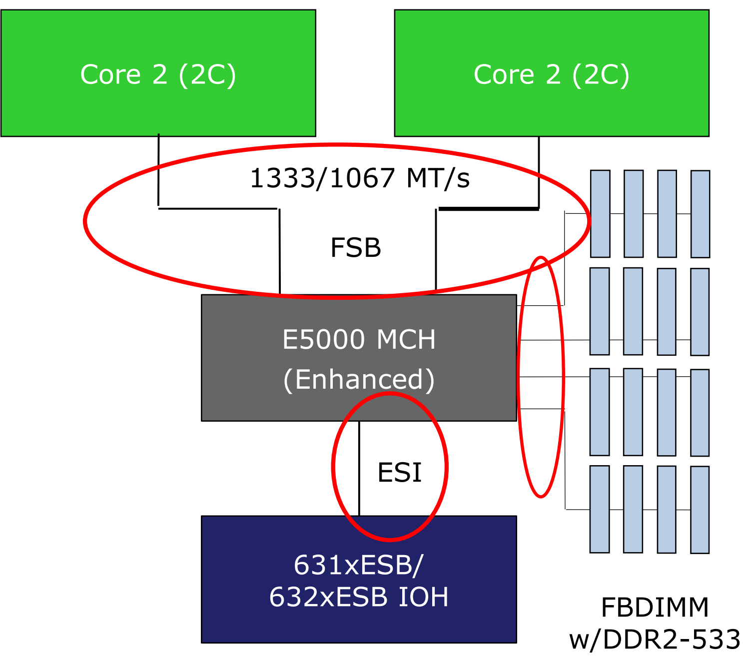
- the processor or processors (in multiprocessors)
- the related chipset as well as
- the interfaces (buses) interconnecting them.

Remark

The designation platform is often understood as the entire system architecture.

*

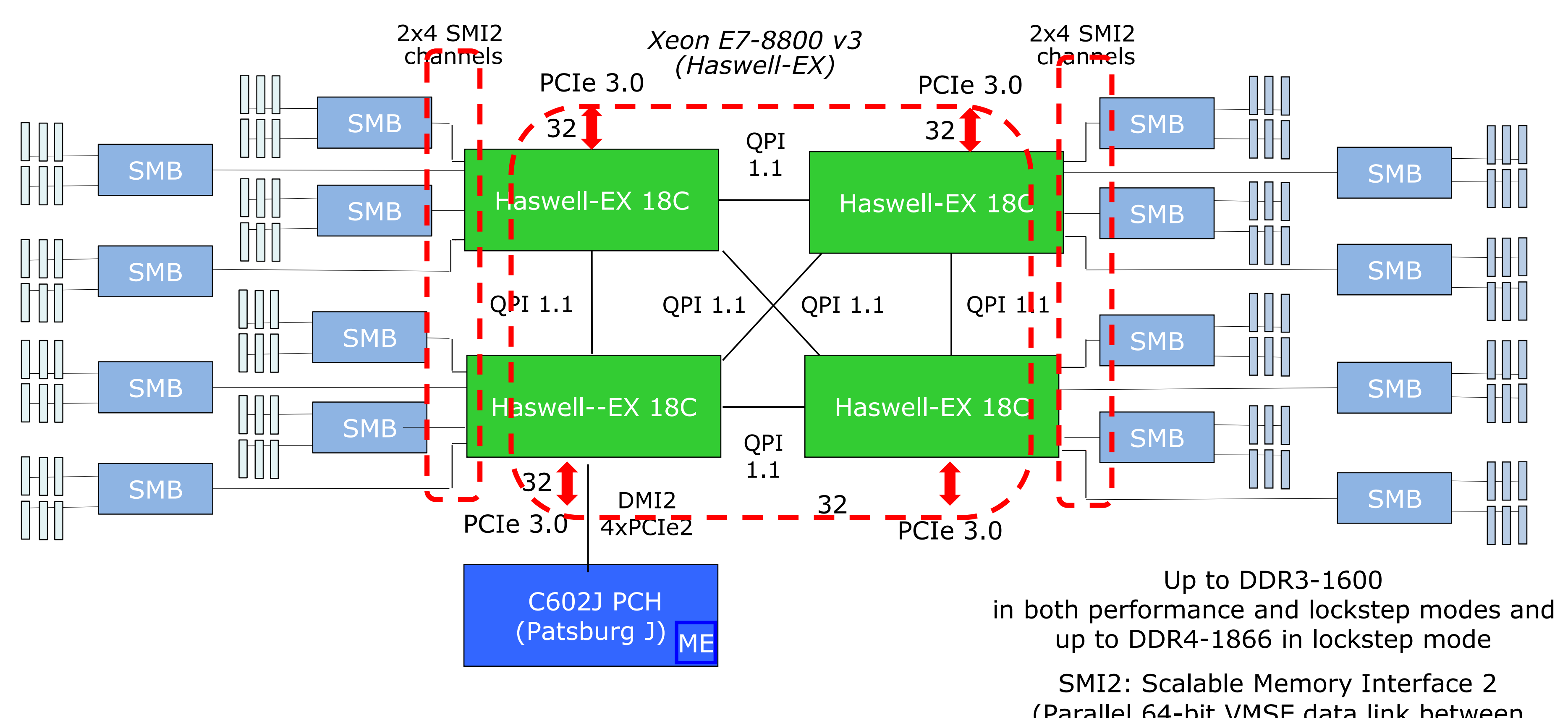
Example 1: A simple traditional 2S server platform



2S: Dual Socket

1.2 The platform concept (3)

Example 2: A recent 4S platform
(The Brickland 4S/8S server platform with Haswell-EX processors (2015)



Up to DDR3-1600
in both performance and lockstep modes and
up to DDR4-1866 in lockstep mode

SMI2: Scalable Memory Interface 2
(Parallel 64-bit VMSE data link between
the processor and the SMB)

QPI 1.1: Up to 9.6 GT/s

SMB: Scalable Memory Buffer
(C112/C114: Jordan Creek 2)
(Performs conversion between the
parallel SMI2 and the parallel
DDR3/DDR4 DIMM interfaces)

ME: Management Engine

C112: 2 DIMMs/channel
C114: 3 DIMMs/channel

Compatibility of platform components (2)

Since platform components are connected via specified interfaces to build an entity, called platform, platform components, such as processors or memories are compatible to a given platform as long as they have the same interfaces (including interface parameters, such as transfer rates etc.).

*

Example: Interface requirements for compatible platform components

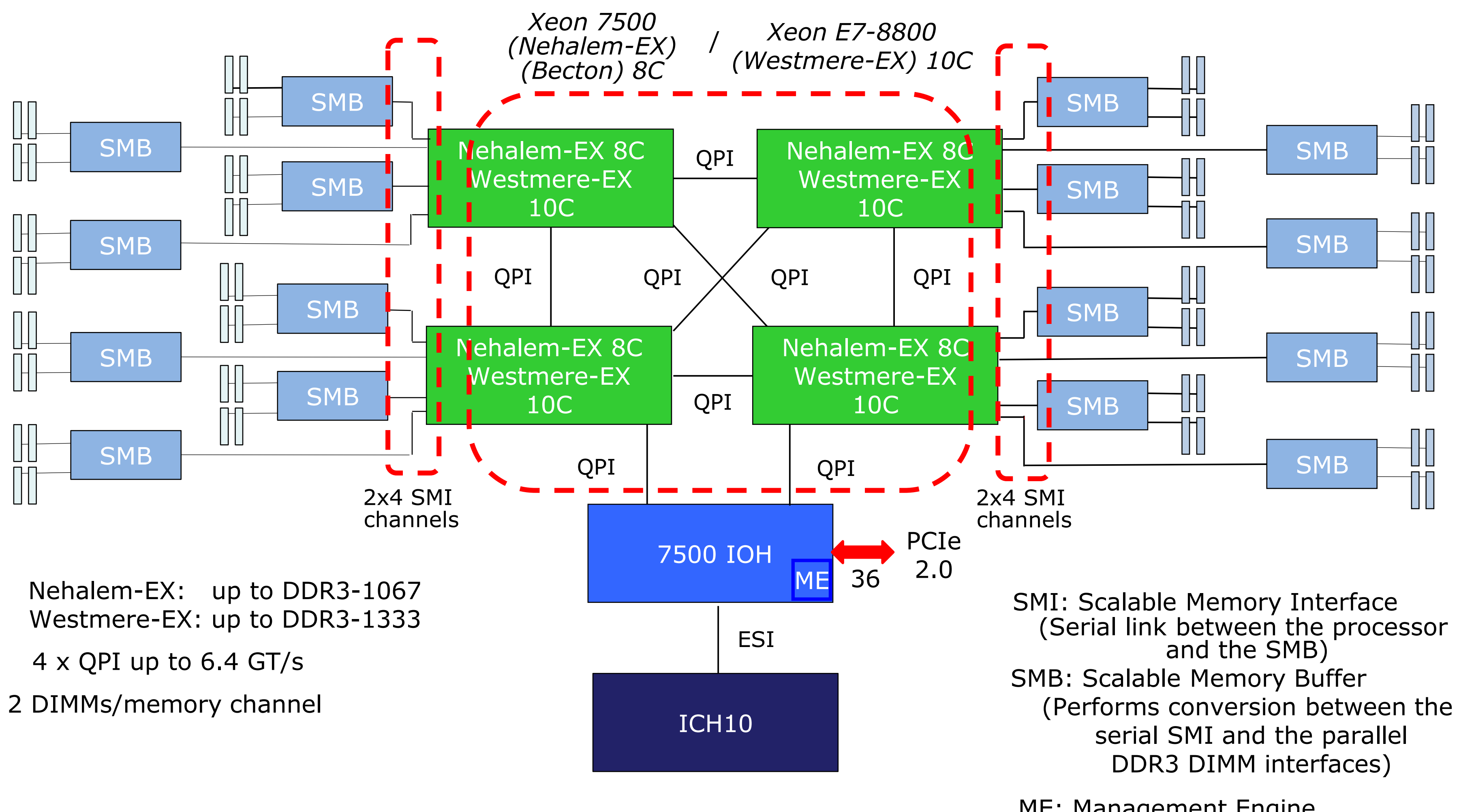


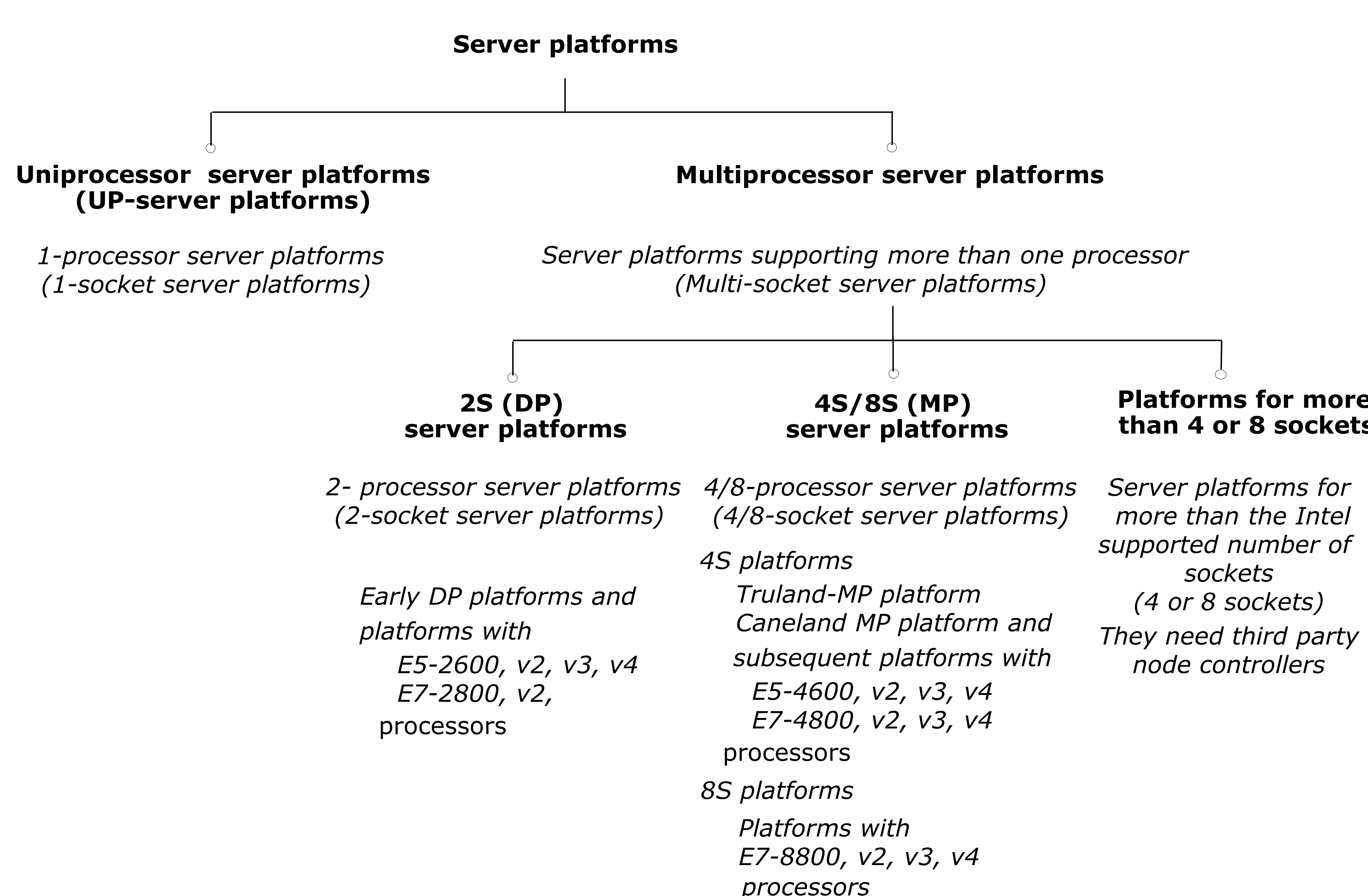
Figure: The Boxboro-EX MP server platform supporting Nehalem-EX/Westmer-EX server processor

Overview of Intel's high-end 4S/8S multicore server platforms and processors

Platform/ Scalability	Core	Techn.	Intro.	High-end 4S/8S server processor lines	Core count	Chipset	Proc. socket
Truland MP (4S)	Pentium 4 MP Prescott	90 nm	3/2005	90 nm Pentium 4 MP (Potomac)	1C	E8500 + ICH5	LGA 604
Truland MP (4S)	Pentium 4 Presc.	90 nm	11/2005	7000 (Paxville MP)	2x1C	E8501 + ICH5	LGA 604
Truland MP (4S)	Pentium 4 Presc.	65 nm	8/2006	7100 (Tulsa)	2x1C	E8501 + ICH5	LGA 604
Caneland (4S)	Core2	65 nm	9/2007	7200 (Tigerton DC) 7300 (Tigerton QC)	2C 2x2C	E7300 (Clarksboro)+ 631x/632x ESB	LGA 604
Caneland (4S)	Penryn	45 nm	9/2008	7400 (Dunnington)	6C	E7300 (Clarksboro)+ 631x/632x ESB	LGA 604
Boxboro-EX (8S)	Nehalem	45 nm	3/2010	7500 (Beckton/ Nehalem-EX)	8C	7500 (Boxboro) + ICH10	LGA 1567
Boxboro-EX (8S)	Westmere	32 nm	4/2011	E7-8800 (Westmere-EX)	10C	7500 (Boxboro) + ICH10	LGA 1567
	Sandy Bidge	32 nm					
Brickland (8S)	Ivy Bridge	22 nm	2/2014	E7-8800 v2 (Ivy Bridge-EX)	15C	C602J (Patsburg J)	LGA 2011-1
Brickland (8S)	Haswell	22 nm	5/2015	E7-8800 v3 (Haswell-EX)	18C	C602J (Patsburg J)	LGA 2011-1
Brickland (8S)	Broadwell	14 nm	6/2016	E7-8800 v4 (Broadwell-EX)	24C	C602J (Patsburg J)	LGA 2011-1
Purley (2S/4S/8S)	Skylake	14 nm	7/2017	Platinum 8100 (Skylake-SP)	28C	C620 (Lewisburg)	LGA 3647

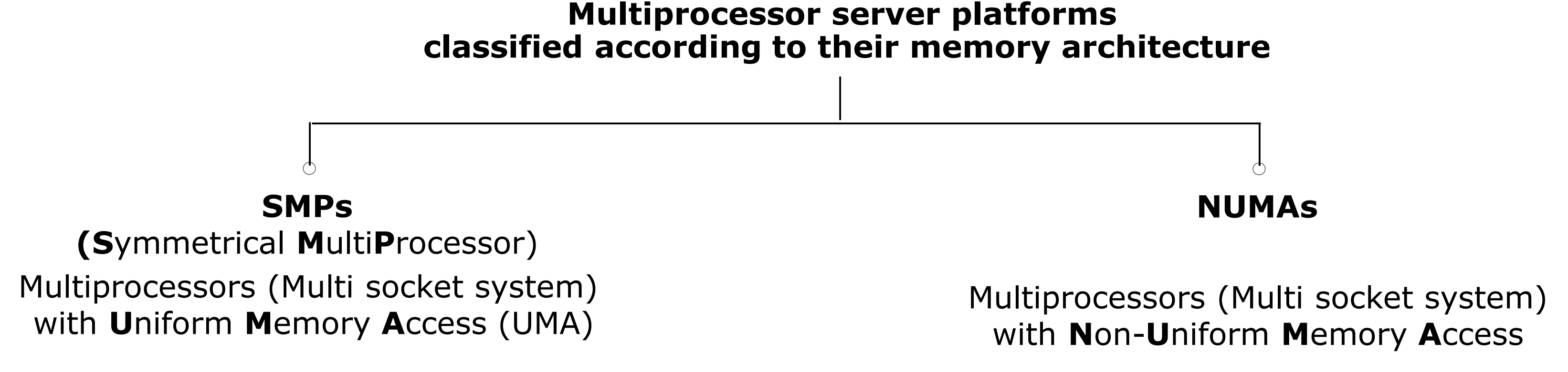
1.3 Server platforms classified according to their scalability

1.3 Server platforms classified according to their scalability



1.4 Multiprocessor server platforms classified according to their memory architecture

1.4 Multiprocessor server platforms classified according to their memory architecture (1)



All processors access main memory by the same mechanism, (e.g. by individual FSBs and an MCH).

Each processor is allocated a part of the main memory (with the related memory space), called the local memory, whereas the rest is considered as the remote memory.

Typical examples

Processor
Processor
FSB
MCH
E.g. DDR2-533
ESI
ICH

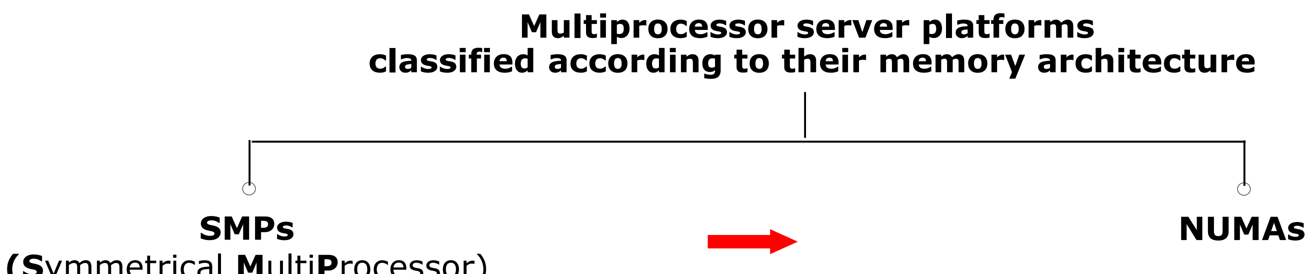
Processor
QPI
Processor
QPI
QPI
IOH[1]
E.g. DDR3-1333
E.g. DDR3-1333
ESI
ICH

[1]ICH: I/O hub

*

ESI: Enterprise System Interface

Main features of SMP-type and NUMA-type server platforms



SMPs (Symmetrical MultiProcessor)

- All processors share the same memory space.
- Consequently, all processors access memory basically with the same latency.

NUMAs

- All processors share the same memory space.
- The local part of the main memory can be accessed by each processor immediately, e.g. via their private on-die memory controller,

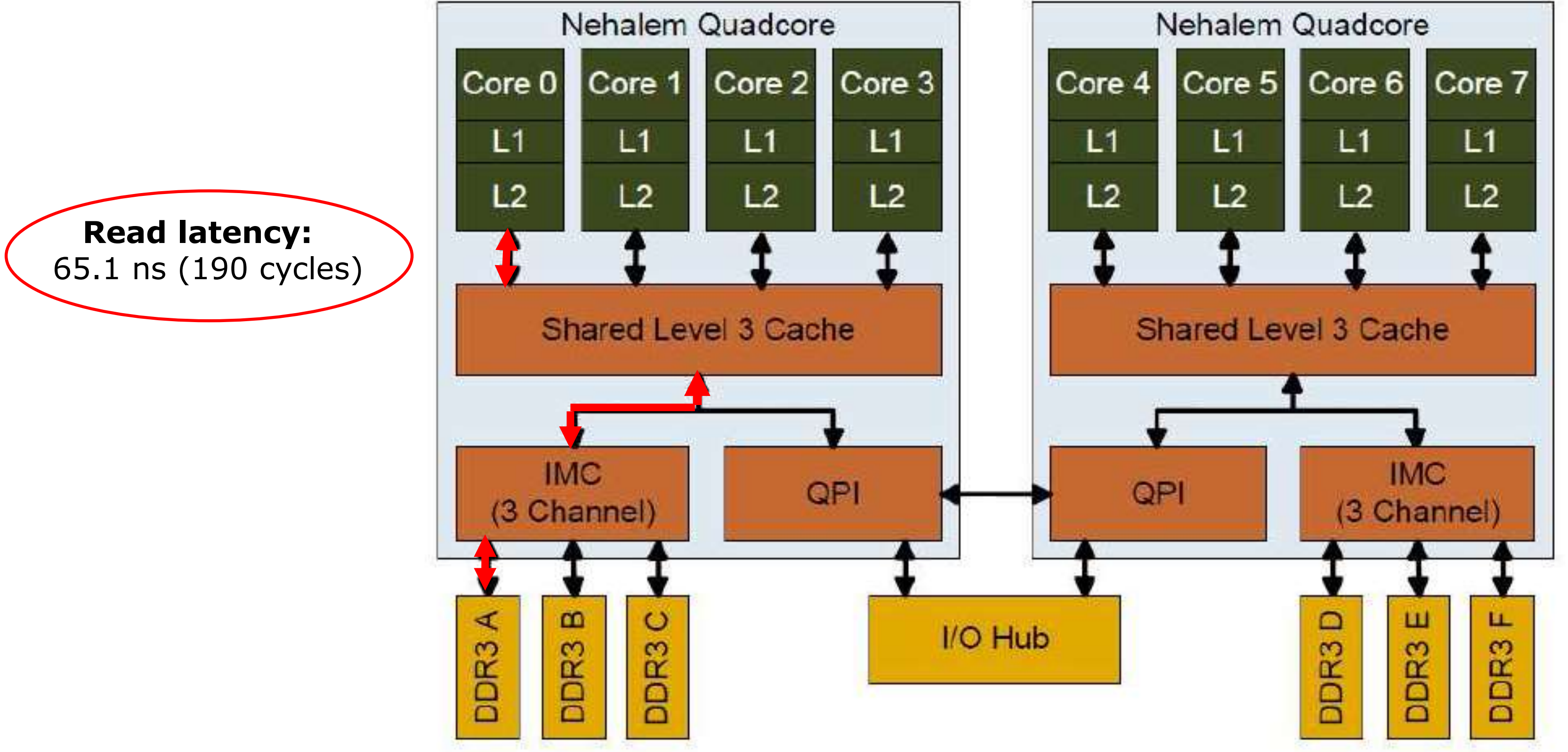
  whereas remote parts of the main memory will be accessed via the processor owing the requested memory part.
- Processors access their local memory space with a significantly shorter latency than their remote memory space, consequently, processors access main memory with different latencies.

*

Example of measured read latencies of uncached data [73]

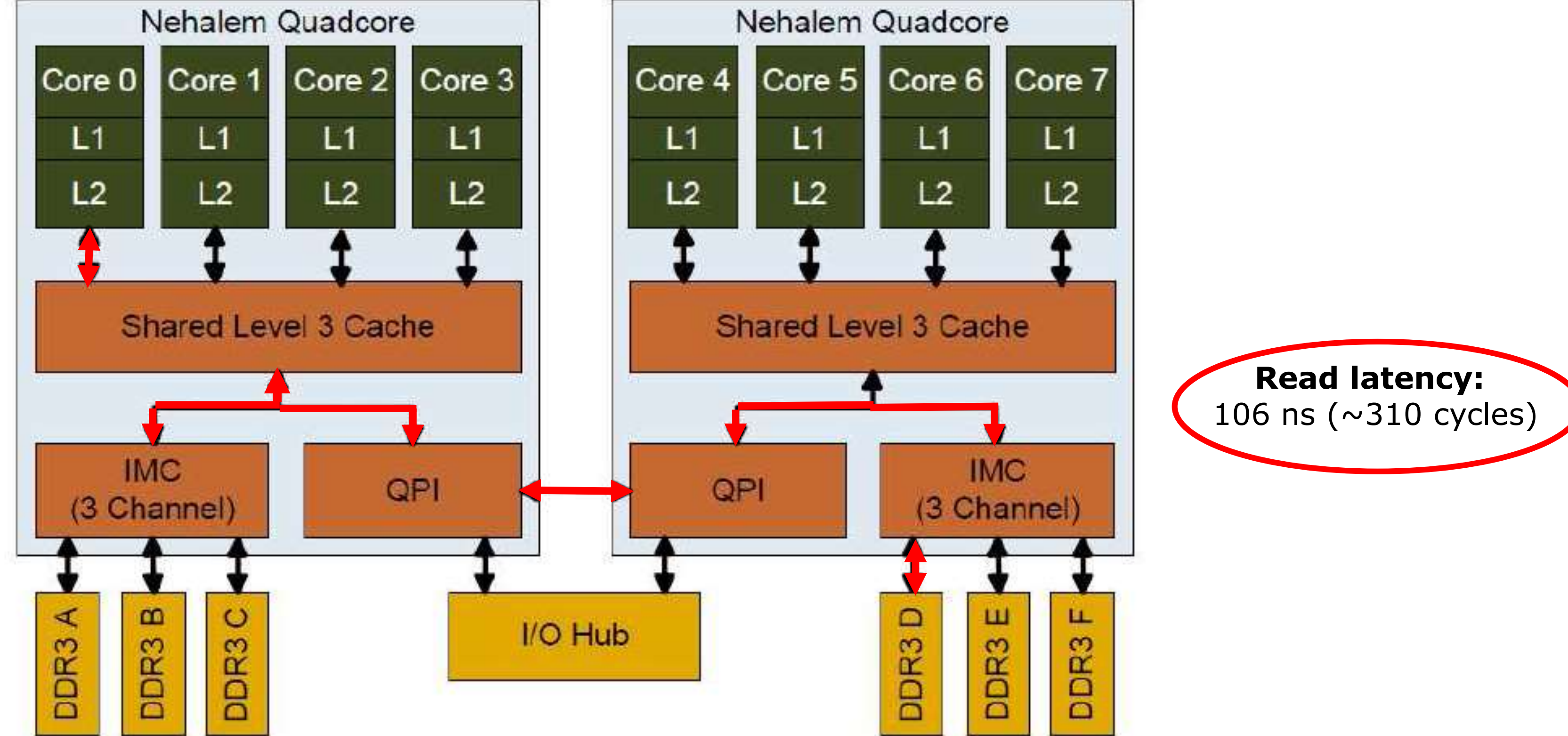
Read latencies depend on whether referenced data is kept in the own or remote main memory space.

a) Read latency of uncached data when referenced data is kept in the own memory space



IMC: Integrated Memory Controller

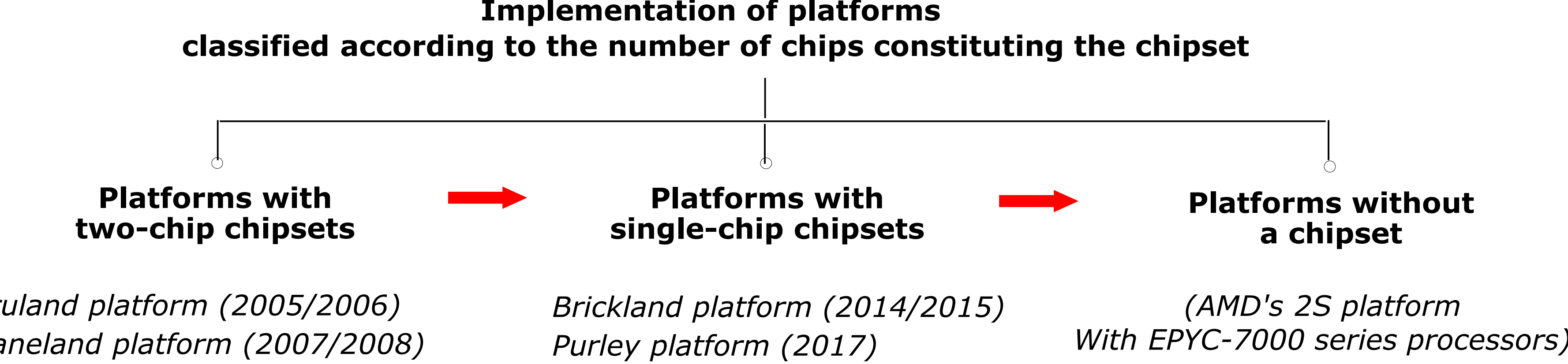
b) Read latency of uncached data when referenced data is kept in the remote memory space [73]



Read latency is increased now by the inter-processor access penalty of 41 ns in this case.

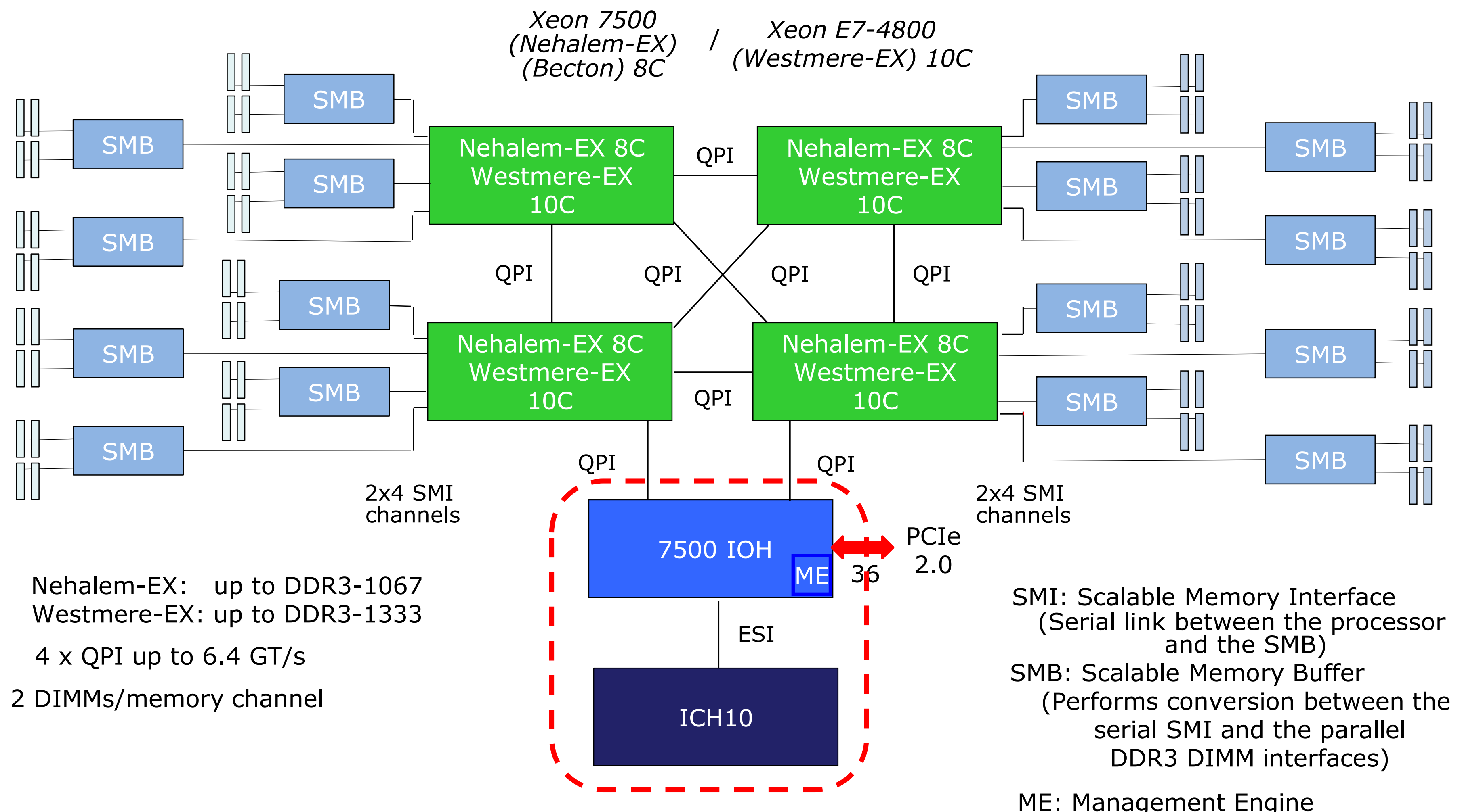
1.5 Platforms classified according to the number of chips constituting the chipset

1.5 Platforms classified according to the number of chips constituting the chipset



*

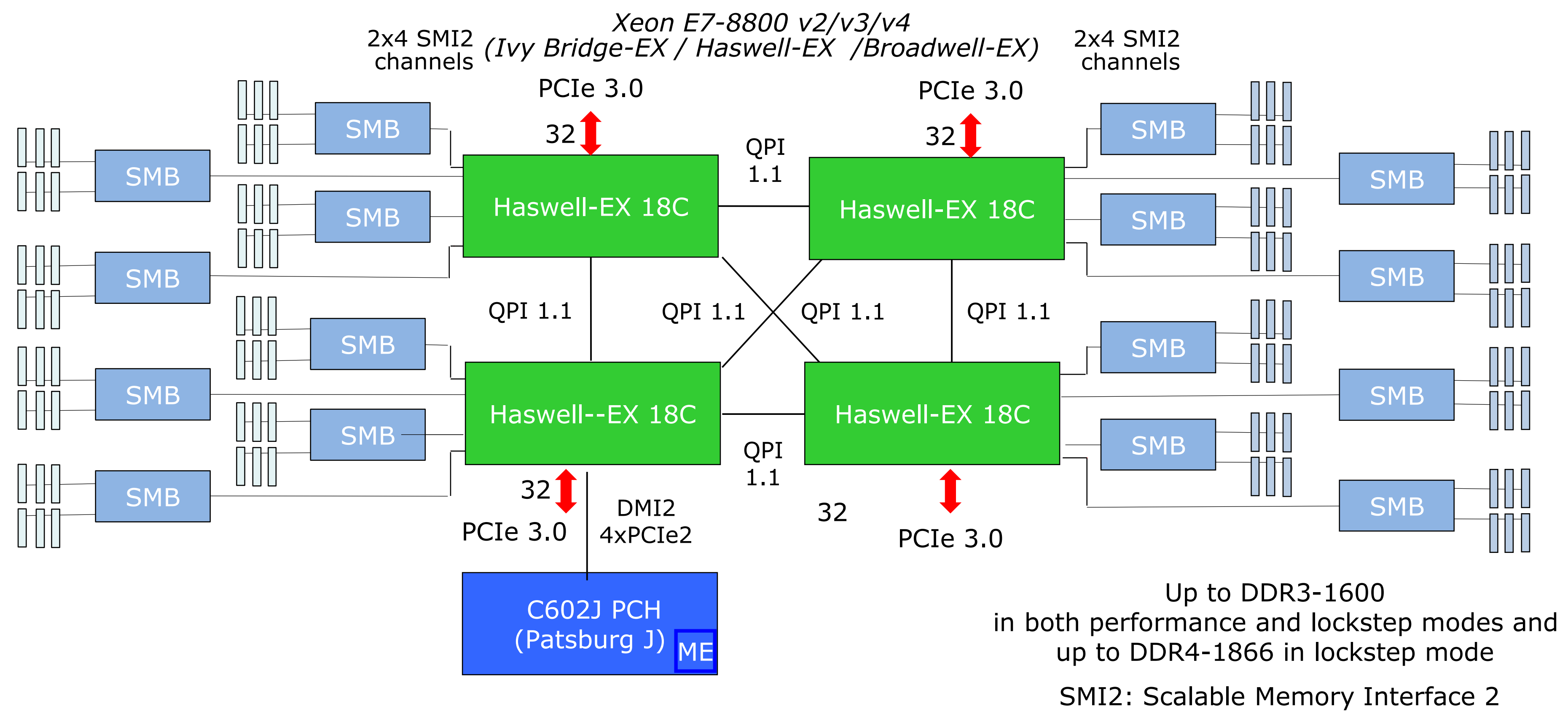
Example of a platform with a 2-chip chipset: The Boxboro-EX 4S/8S server platform supporting Nehalem-EX/Westmere-EX processors



Nehalem-EX aimed Boxboro-EX MP server platform (for up to 10 C)

1.5 Platforms classified according to the number of chips constituting the chipset (3)

Example of a platform with a single-chip chipset: The Brickland-EX 4S/8S server platform supporting Ivy Bridge-EX/Haswell-EX/Broadwell-EX processors



Up to DDR3-1600
in both performance and lockstep modes and
up to DDR4-1866 in lockstep mode

SMI2: Scalable Memory Interface 2
(Parallel 64-bit VMSE data link between the processor and the SMB)

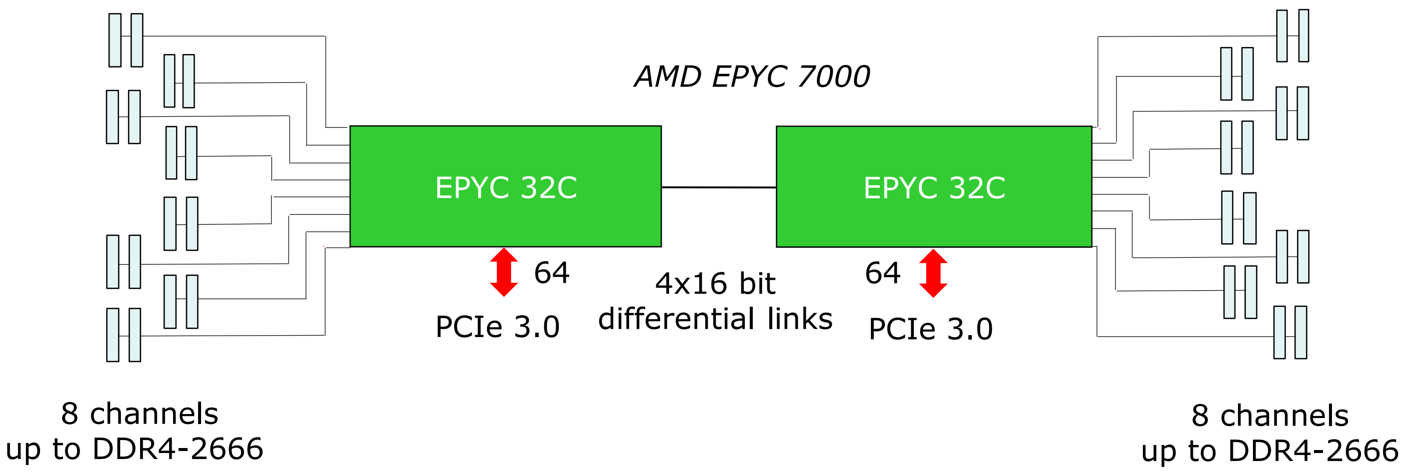
SMB: Scalable Memory Buffer
(C112/C114: Jordan Creek 2)
(Performs conversion between the parallel SMI2 and the parallel DDR3/DDR4 DIMM interfaces)

C112: 2 DIMMs/channel
C114: 3 DIMMs/channel

QPI 1.1: Up to 9.6 GT/s

ME: Management Engine

Example of a platform without a chipset: AMD's 2S server platform supporting EPYC 7000 processors



1.6 Naming schemes of Intel's servers

1.6 Naming schemes of Intel's server processors

a) Intel's naming scheme for servers until 2005

Until 2005 Intel named their servers by identifying clock frequency and server configuration, like Xeon 2.8 GHz DP or Xeon 2.0 GHz MP, since Intel's IC technology was superior to their competitors and provided higher clock frequencies.

Remark

- At that time AMD manufactured their processors in their own foundry, nevertheless by a less advanced IC technology, so AMD's processors had lower clock frequencies vs. Intel's processors.
- To hide this deficiency, AMD introduced a new naming scheme for their processors beginning with their Athlon XP desktop line in 2003.
- The new naming scheme consisted of a four digit number and the +character., e.g. Athlon 1600+.

  AMD interpreted it as the relative performance level related to a given AMD DT model (Athlon 1.4 GHz).

  The relative performance was calculated as the average value of a wide range of benchmarks.

  The actual clock frequencies however, were lower figures, e.g. the Athlon XP 1600+ had a clock frequency of 1.4 GHz.

b) Intel's naming scheme for servers used between 2005 and 2011

- In the first years of the 2000's clock frequencies became stagnating, as indicated below [168].

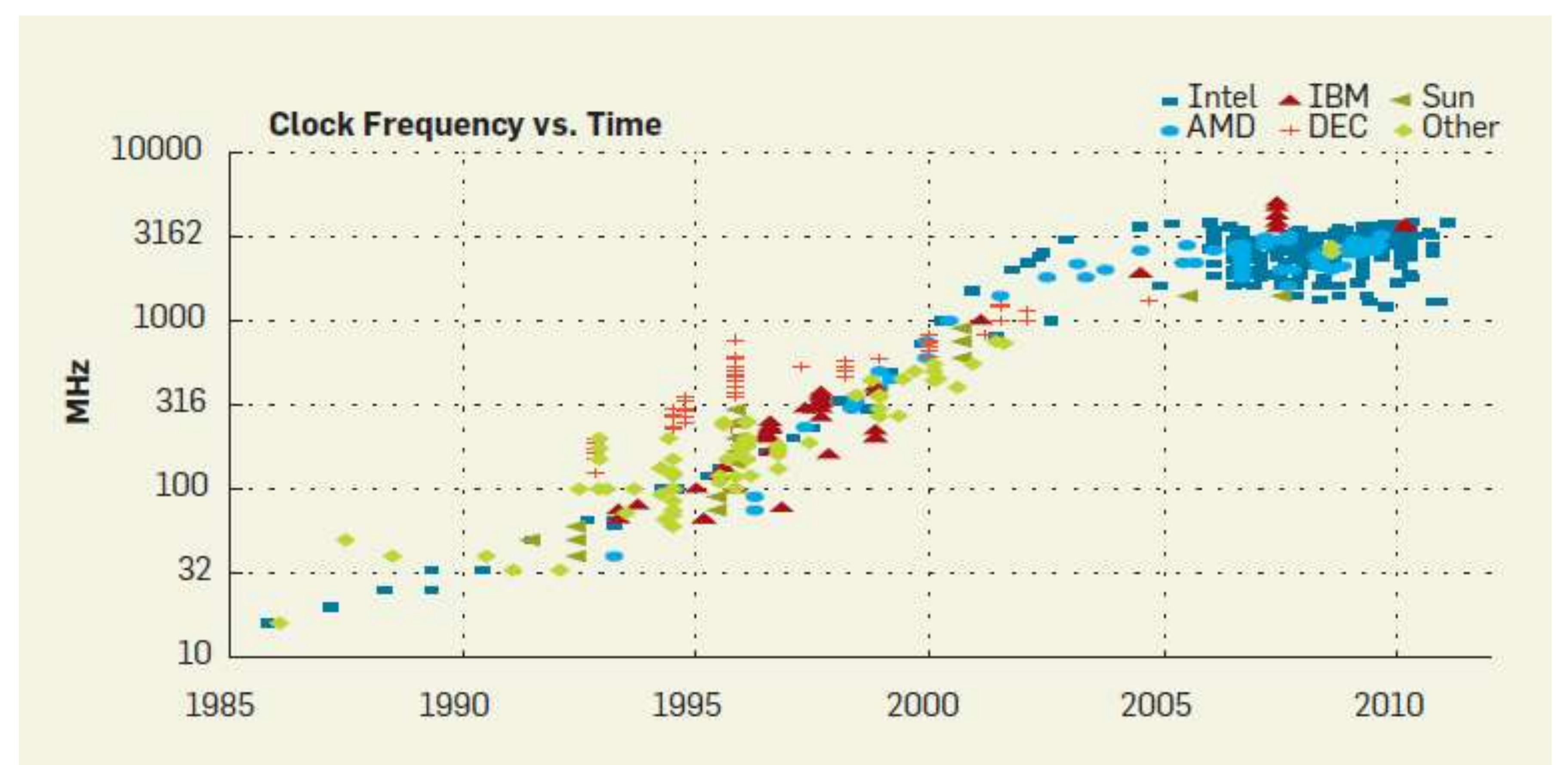


Figure: Historical growth of clock rates [168]

b) Intel's naming scheme for servers used between 2005 and 2011 -2

- As an aftermath Intel had to disrupt using its previous naming scheme and introduced an AMD like naming scheme in 2005, as follows:

Accordingly, Intel's subsequent MP (4S) server processor lines were designated as follows:

Line	Processor	Based on	Intro
7000	Paxville MP	Pentium 4 Prescott MP	2005
7100	Tulsa	Pentium 4 Prescott MP	2006
7200	Tigerton DC	Core 2	2007
7300	Tigerton QC	Core 2	2007
7400	Dunnington	Penryn	2008
7500	Beckton	Nehalem	2010

The drawback of this naming scheme is that it does not reveal any significant features of the related server families.

c) Intel's renewed naming scheme for servers introduced with the Westmere line in -1

- In 4/2011 Intel renewed its entire naming scheme to reflect more details in the designations.
- This resulted in the following new naming scheme for servers:

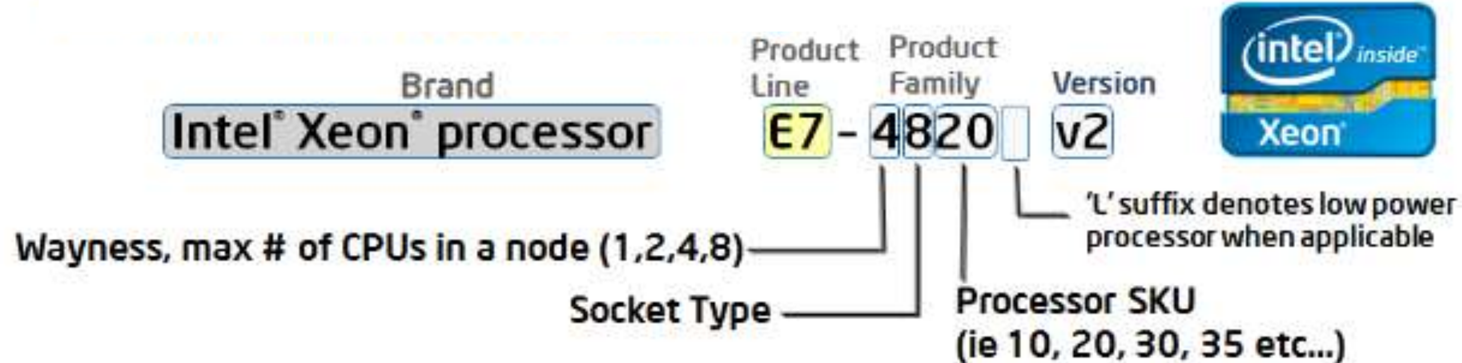


Figure: Intel's new Xeon naming scheme [127]

- Interpretations
  - Product line: E3, E5 or E7 (E: Enterprise).
  - Wayness: How many processors are natively supported in a system
  - Socket type: Signifies processor performance,
    - 8: high performance
    - 6: effective performance
    - 4: entry level
  - Processor SKU: Last two digits of the model number (SKU: Storage Keeping Unit)
  - Version: The same version numbers indicate a common microarchitecture
    - Westmere/Sandy Bridge (without any version number)
    - v2: Ivy Bridge
    - v3: Haswell
    - v4: Broadwell etc.

*

1.6 Naming schemes of Intel’s servers (5)

Remarks

1) The designation of the server generations v2, v3 etc. does not coincides with the designation of the microarchitecture generations, as shown below:

Microarchitecture	Architecture generation	Server generation
Westmere		(without designation) only E7
Sandy Bridge	Gen. 2	(without designation) only E3/E5
Ivy Bridge	Gen. 3	v2
Haswell	Gen. 4	v3
Broadwell	Gen. 5	v4

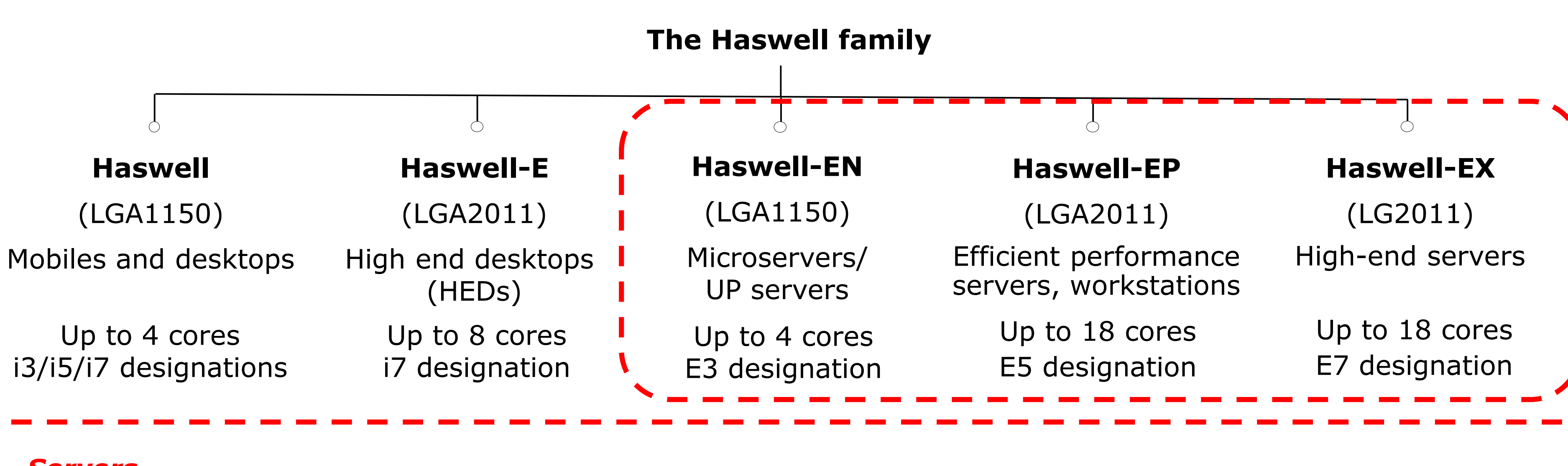
2) The Itanium line has not been renamed, it remained the 9000-line, with model numbers of 9300, 9500, up to 9700

Intel's renewed naming scheme for servers introduced with the Westmere line -2

Accordingly, Intel high-end (EX) server lines became designated as follows:

High-end server lines	Models	Scalability	Core count up to	Socket	Intro
Westmere-EX	E7-2800 E7-4800 E7-8800	2-socket 4-socket 8-socket	10C	LGA 1567	2011
Ivy Bridge-EX	E7-28xx v2 E7-48xx v2 E7-88xx v2	2-socket 4-socket 8-socket	15C	LGA 2011-1	2014
Haswell-EX	E7-48xx v3 E7-88xx v3	4-socket 8-socket	14C 18C	LGA 2011-1	2015
Broadwell-EX	E7-48xx v4 E7-88xx v4	4-socket 8-socket	16C 22C	LGA 2011-1	2016

Example: The full range of Intel's server product lines in the Haswell family



Servers

Microservers

E3-1275L/1265L v3, 4C+G, HT, 6/2013 and 5/2014
E3-1240L/1230L/1220L v3, 2C/4C, HT, 6/2013 and 5/2014

UP Servers

E3-12x5/12x6 v3, 4C+G, HT, 6/2013 and 5/2014
E3-12x0/12x1 v3, 4C, HT, 6/2013 and 5/2014

Workstations

E5-16xx v3, 4/6/8C, 9/2014

E5-24xx v3, 4/6/8/10C, 1/2015

2S-Servers

E5-26xx v3, 4/6/8/10/12/14/16/18C, 9/2014

E5-46xx v3, 6/10/12/14/16/18C, 6/2015

4S/8S-Servers

E7-48xx v3, 8/10/12/14C, 5/2015
E7-88xx v3, 4/10/16/18C, 5/2015

Drawback of the implementation of Intel's - Westmere to Broadwell based - 2S/4S/8S servers

There is a large number of different implementation alternatives of Intel's Westmere to Broadwell based 2S to 8S server processors, as the next Figure indicates.

*

1.6 Naming schemes of Intel's servers (9)

Example: E5/E7 platform options built up of Ivy Bridge based server processors [117]

Romley platform

EN: Entry level

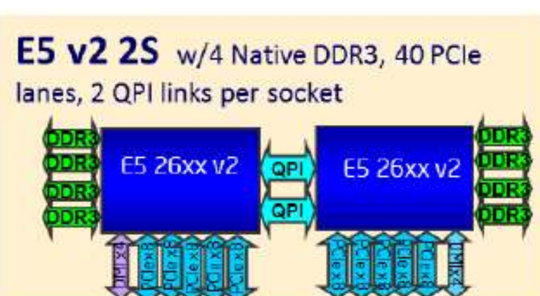
EP: Efficient Performance level options

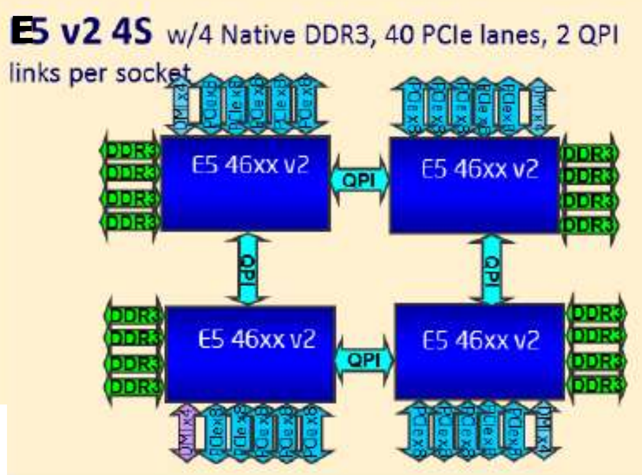
E5 v2 2S w/3 Native DDR3, 20 PCIe lanes, 1 QPI link per socket

E5-24xx v2 | QPI | E5-24xx v2

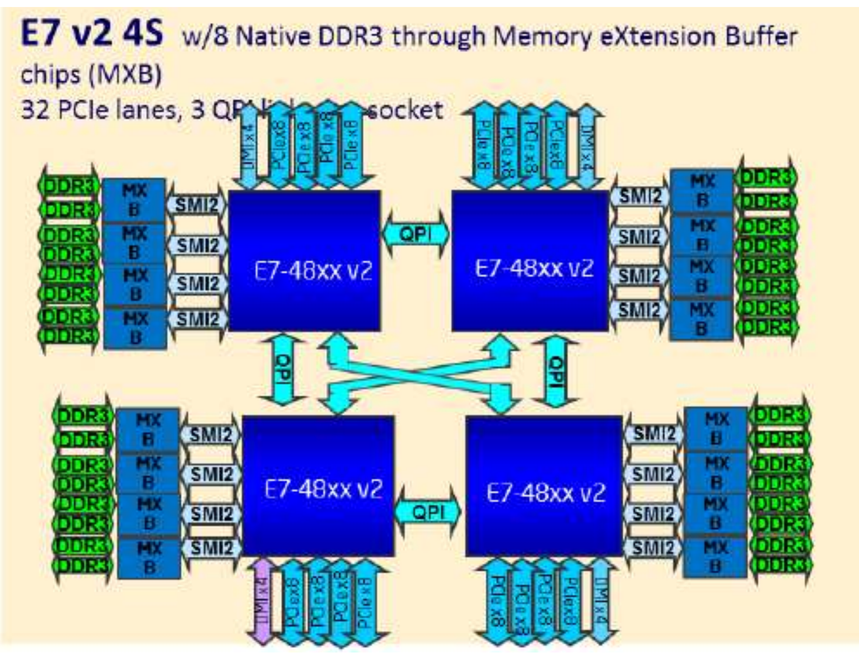
DDR3 | DDR3 | DDR3

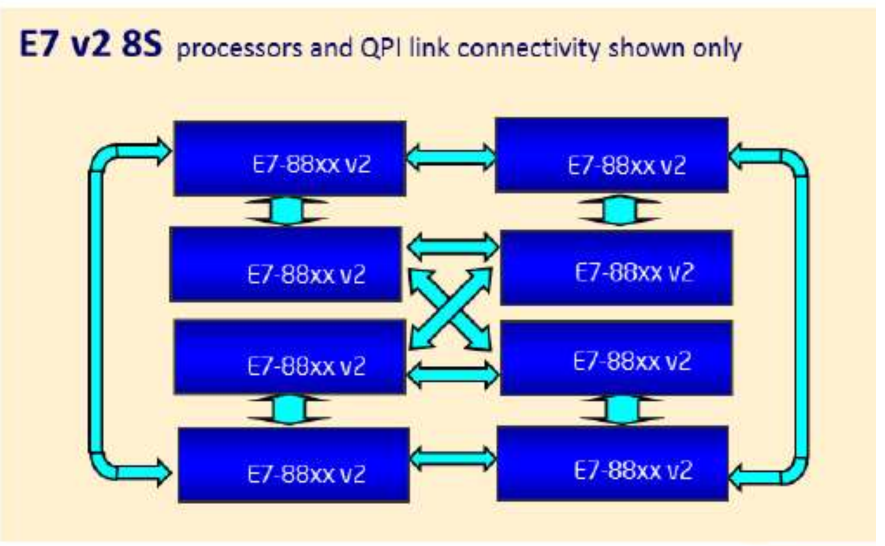
DMI x2 | PCIe x8 | PCIe x8 | PCIe x4





Brickland platform: High-end (EX) platform options





Example: E5/E7 processor features of Intel's Ivy Bridge based server processors [117]

The main differences are between implementing the E5-EN, E5-EP and the E7-EX Ivy Bridge based server processors, as follows:

	E5-EN lines	E5- EP lines	E7-EX lines
Type of memory attachment to the on-die MC(s)	Directly attached DDR3 memory channels to a single or dual MCs		Attachment via 4 low line count proprietary 64-bit parallel channels (SMI2 channels) with on-board memory buffers, while 2 DDR3/4 channels are connected to each mem. buffer
No. of DDR3 memory channels	3 DDR3 channels	4 DDR3 channels	8 DDR3 channels
No. of QPI links	1 QPI link	2 QPI links	3 QPI links
No. of PCIe lanes	3x8 lanes	5x8 lanes	4x8 lanes
PCH	2S options: C600 (Patsburg) 4S options: C610 (Wellsburg)		C602 J (Patsburg J)

d) Intel's latest naming scheme for servers introduced with the Skylake-SP line in 2017

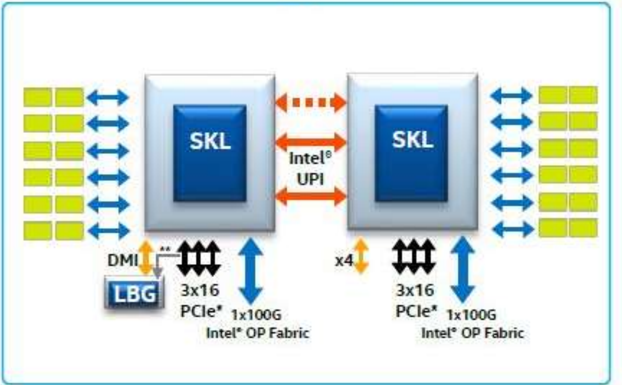
It aims at

- unifying the processor implementations of the 2S, 4S, 8S configurations,
- introducing four performance grades instead of two (E7/E5) and

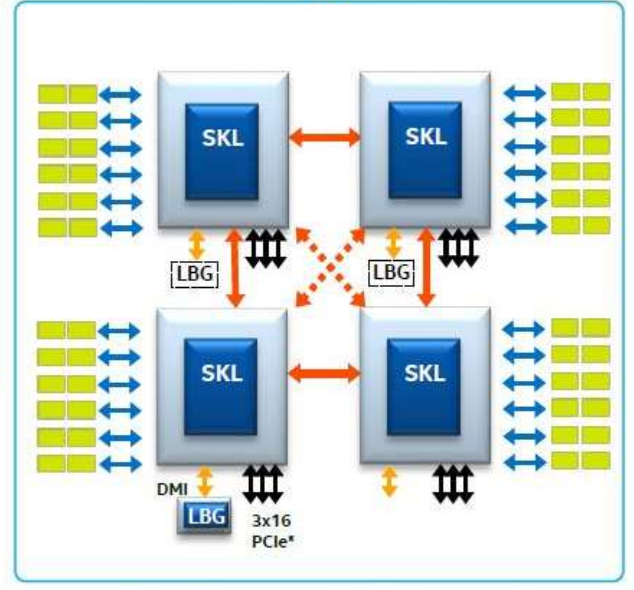
as indicated in the following figures.

*

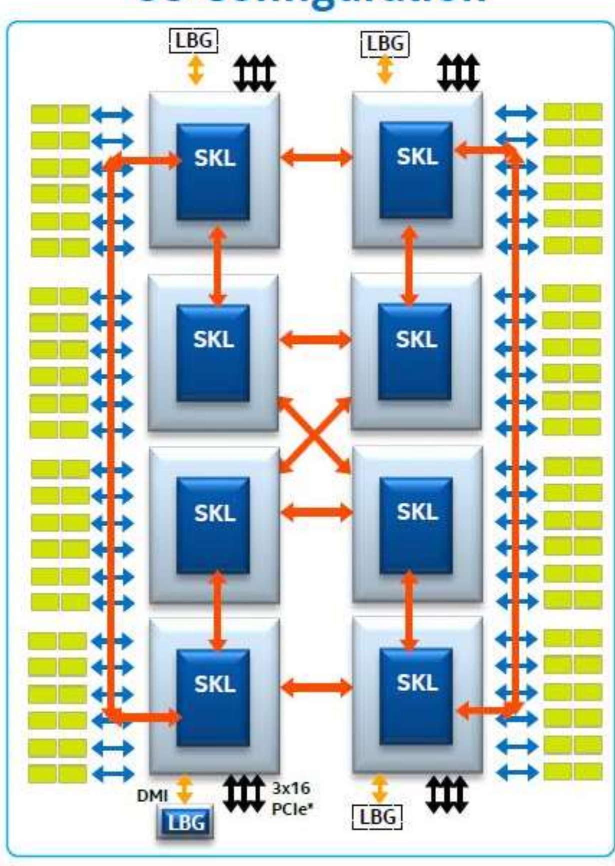
Unifying the processor implementation of the 2S to 8S configurations [139]



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

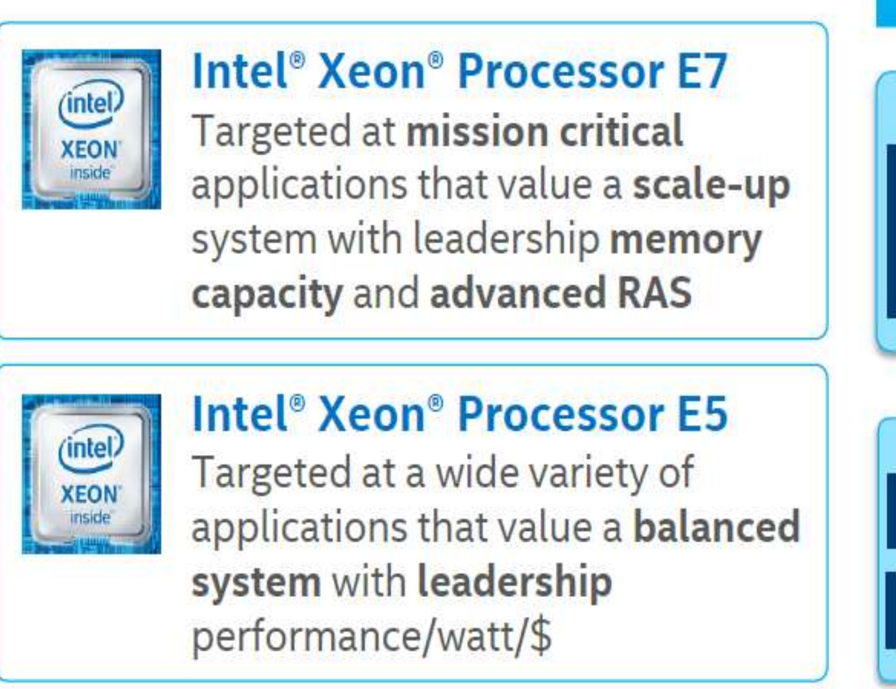


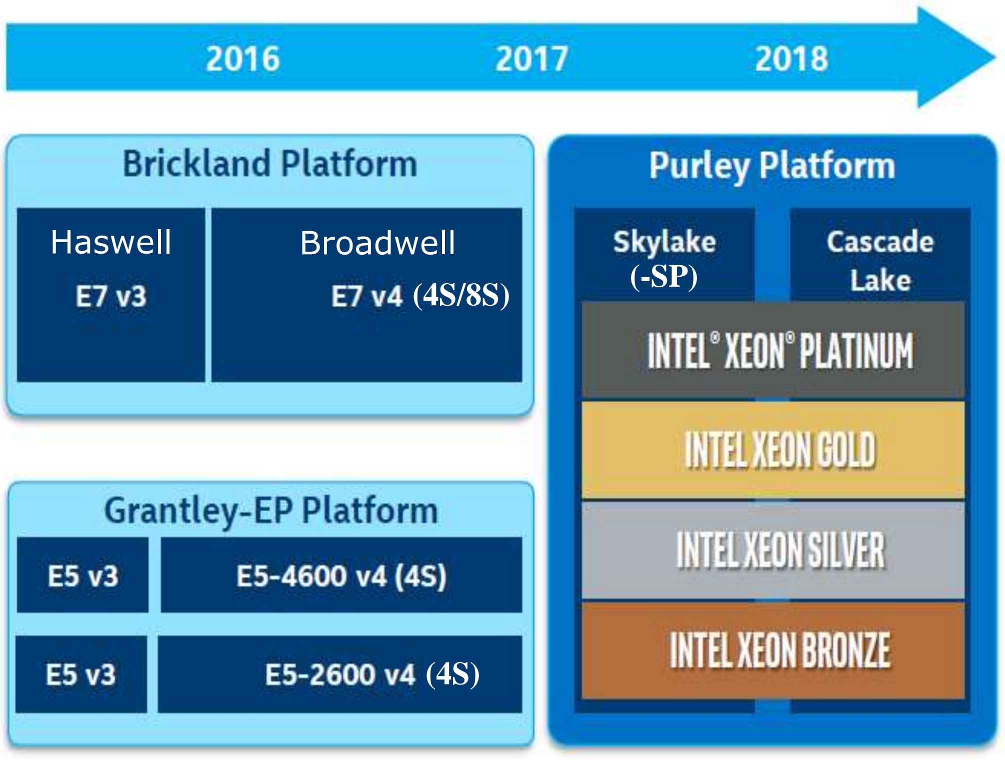
(4S-2UPI & 4S-3UPI shown)



LBG: Lewisburg (PCH)
OP: OmniPath

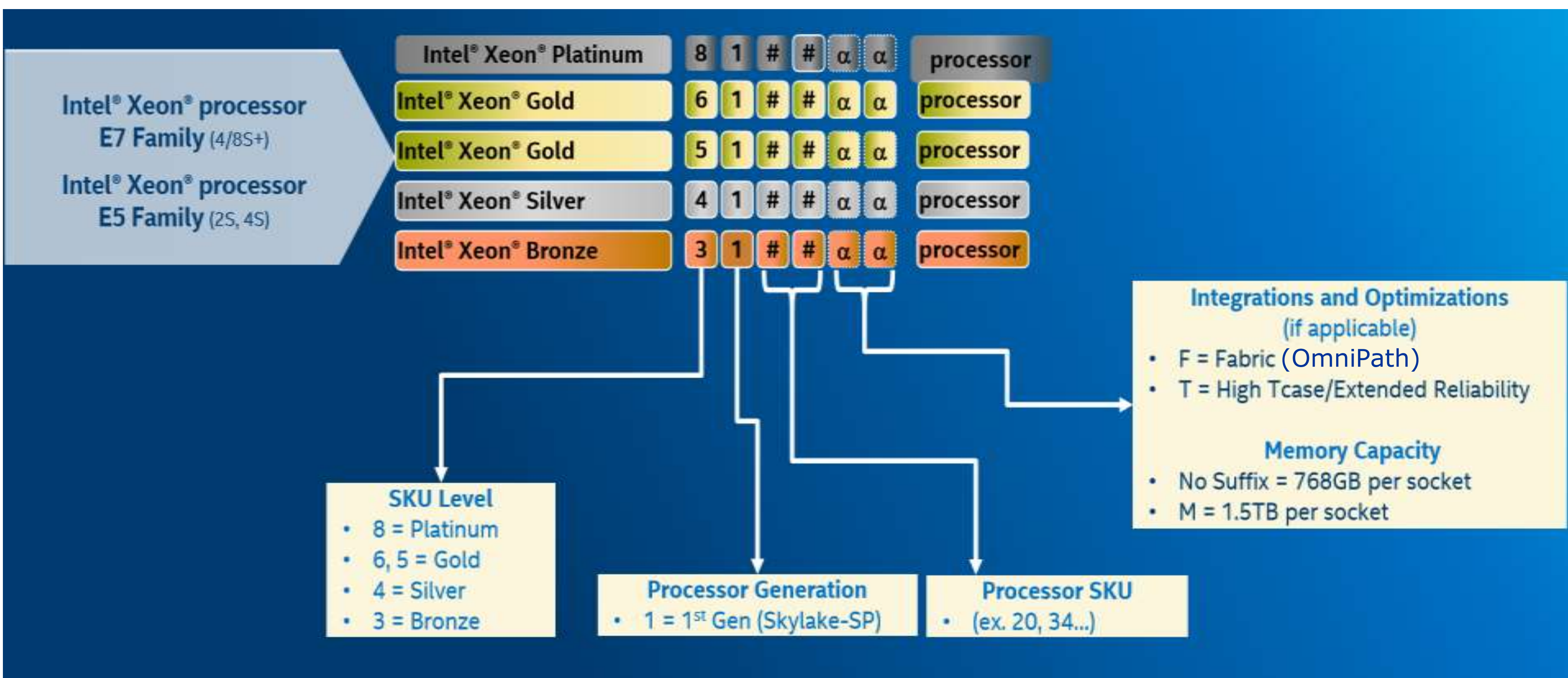
Introducing four performance grades instead of two (E7/E5)[139]





CONVERGED PLATFORM WITH INNOVATIVE SKYLAKE-SP MICROARCHITECTURE

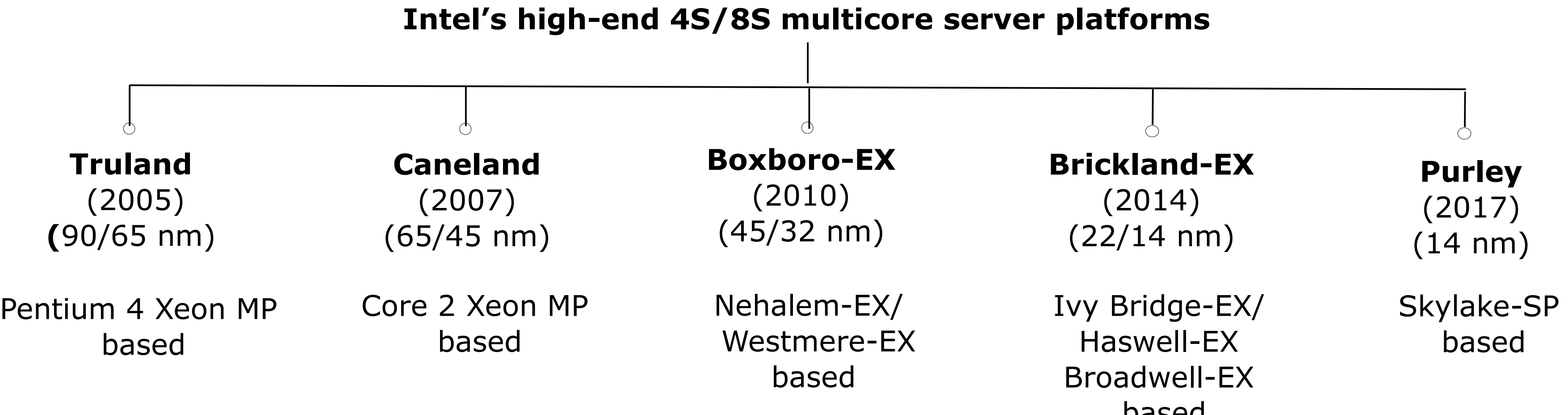
Corresponding new model naming starting with the Skylake-SP processor line [140]



*

1.7 Overview of Intel's 4S/8S high-end multicore server platforms

1.7 Overview of Intel's high-end 4S/8S multicore server platforms -1

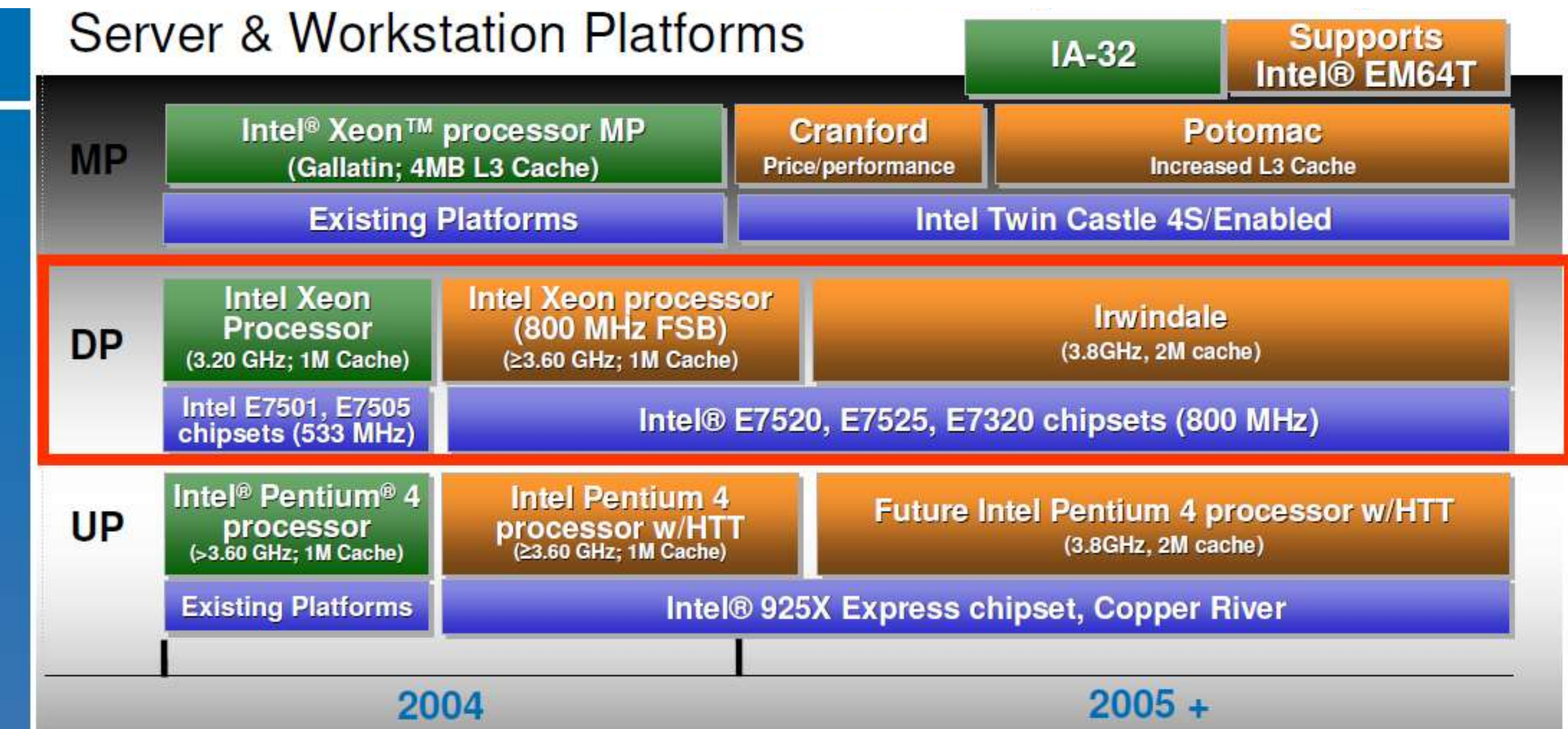


*

Overview of Intel's high-end multicore 4S/8S platforms and processors

Platform	Core	Techn.	Intro.	High-end server processor lines	Core count	Chipset	Proc. socket
Truland MP	Pentium 4 MP Prescott	90 nm	3/2005	90 nm Pentium 4 MP (Potomac)	1C	E8500 + ICH5	LGA 604
	Pentium 4 Presc.	90 nm	11/2005	7000 (Paxville MP)	2x1C	E8501 + ICH5	
	Pentium 4 Presc.	65 nm	8/2006	7100 (Tulsa)	2x1C		
Caneland MP	Core2	65 nm	9/2007	7200 (Tigerton DC) 7300 (Tigerton QC)	2C 2x2C	E7300 (Clarksboro)+ 631x/632x ESB	LGA 604
	Penryn	45 nm	9/2008	7400 (Dunnington)	6C		
Boxboro-EX	Nehalem	45 nm	3/2010	7500 (Beckton/ Nehalem-EX)	8C	7500 (Boxboro) + ICH10	LGA 1567
	Westmere	32 nm	4/2011	E7-4800 (Westmere-EX)	10C		
	Sandy Bidge	32 nm					
Brickland	Ivy Bridge	22 nm	2/2014	E7-4800 v2 (Ivy Bridge-EX)	15C	C602J (Patsburg J)	LGA 2011-1
	Haswell	22 nm	5/2015	E7-4800 v3 (Haswell-EX)	14C		
	Broadwell	14 nm	6/2016	E7-4800 v4 (Broadwell-EX)	16C		
Purley	Skylake	14 nm	7/2017	Platinum 8100 (Skylake-SP)	28C	C620 (Lewisburg)	LGA 3647

Remark: Remember of Intel's transfer to 64-bit ISA in their server lines [97]



*

2. Evolution of Intel's high-end multicore 4S/8S server platforms

- 2.1 Evolution of key features of Intel's high-end 4S/8S multicore server platforms
- 2.2 The memory bandwidth bottleneck
- 2.3 Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor
- 2.4 The basic design space of connecting memory subsystems to server platforms
- 2.5 The FSB bottleneck in FSB configurations and its resolution
- 2.6 Overview of the evolution of Intel's high-end 4S/8S server platforms

2.1 Evolution of key features of Intel's high-end 4S/8S multicore server platforms

2.1 Evolution of key features of Intel's high-end 4S/8S multicore server platforms

Intel's high-end multicore server processors had supported first 4-socket (4S or MP) and later 8-socket (8S) platforms, as indicated in the Table below.

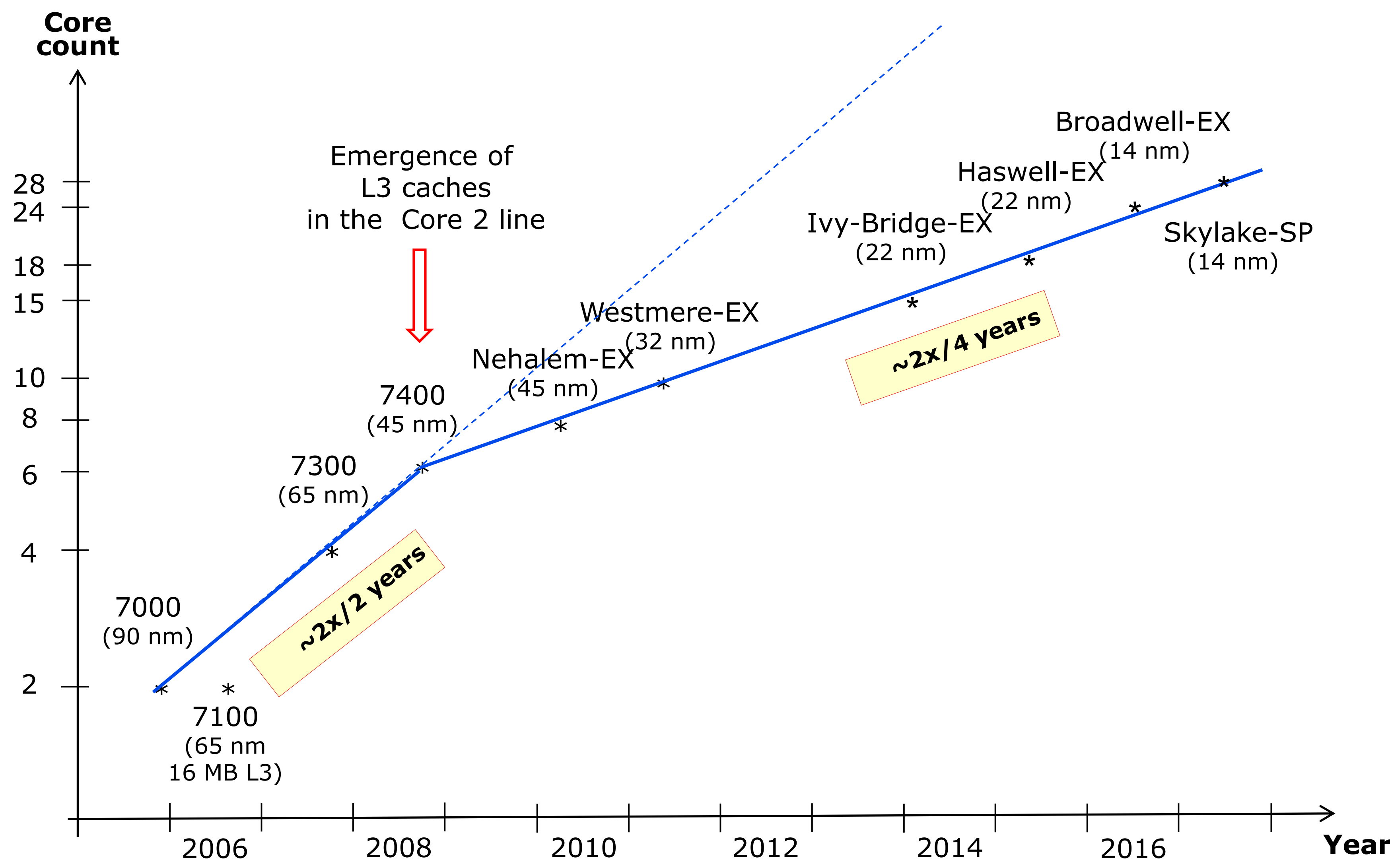
Platform	Core	Techn.	Intro.	High-end 4S/8S server processor lines	Core count	Scalability
Truland MP	Pentium 4 MP Prescott	90 nm	3/2005	90 nm Pentium 4 MP (Potomac)	1C	4-socket servers (4S)
	Pentium 4 Presc.	90 nm	11/2005	7000 (Paxville MP)	2x1C	
	Pentium 4 Presc.	65 nm	8/2006	7100 (Tulsa)	2x1C	
Caneland	Core2	65 nm	9/2007	7200 (Tigerton DC) 7300 (Tigerton QC)	2C 2x2C	4-socket servers (4S)
	Penryn	45 nm	9/2008	7400 (Dunnington)	6C	
Boxboro-EX	Nehalem	45 nm	3/2010	7500 (Beckton/ Nehalem-EX)	8C	8-socket servers (8S)
	Westmere	32 nm	4/2011	E7-8800 (Westmere-EX)	10C	
	Sandy Bidge	32 nm				
Brickland	Ivy Bridge	22 nm	2/2014	E7-8800 v2 (Ivy Bridge-EX)	15C	8-socket servers (8S)
	Haswell	22 nm	5/2015	E7-8800 v3 (Haswell-EX)	18C	
	Broadwell	14 nm	6/2016	E7-8800 v4 (Broadwell-EX)	24C	
Purley	Skylake	14 nm	7/2017	Platinum 8100 (Skylake-SP)	28C	8-socket servers (8S)

Table: Intel's high-end 4S/8S multicore server platforms and processors

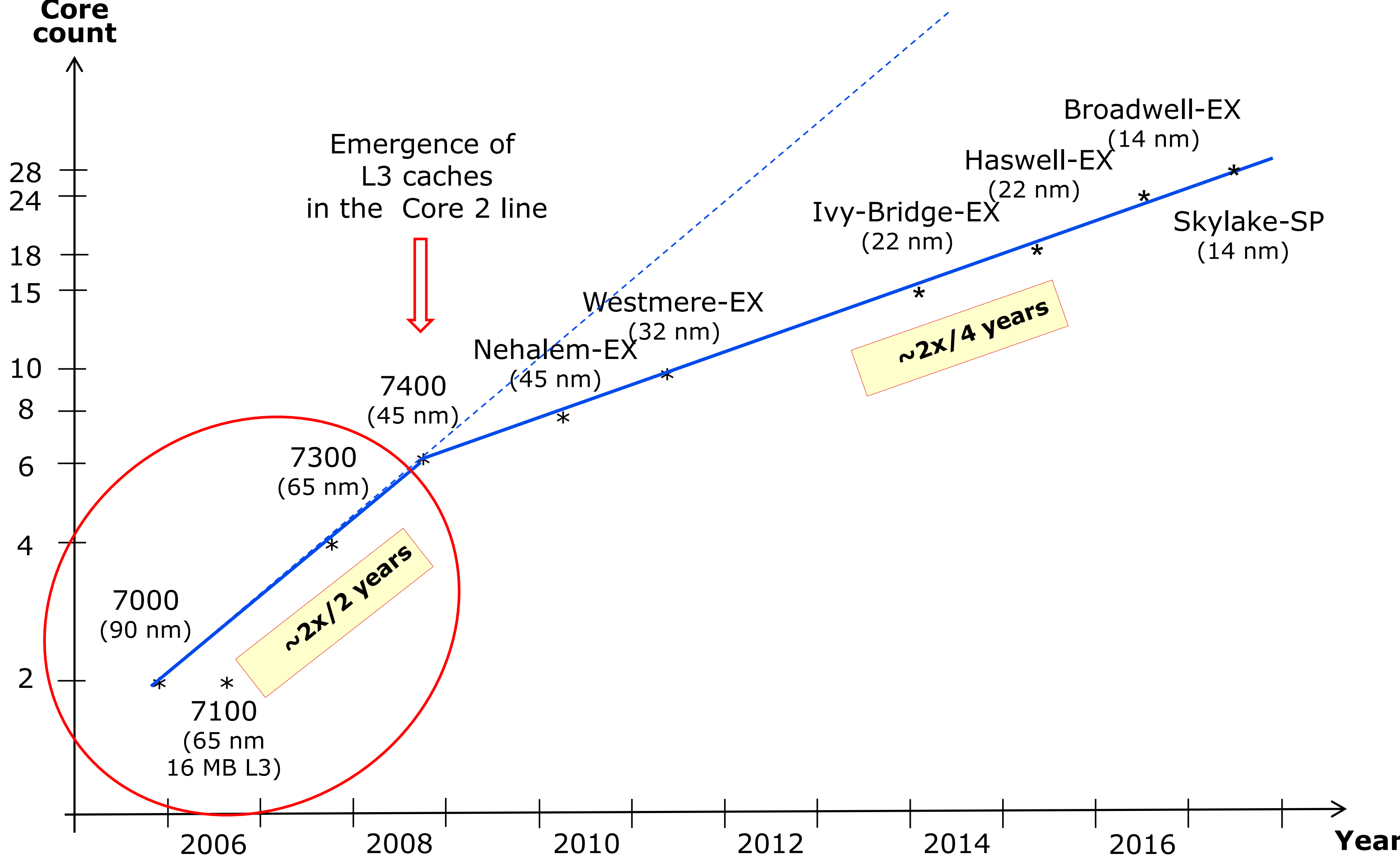
a) Evolution of core counts in Intel's high-end 4S/8S multicore server processors -1

	Core	Techn.	Intro.	High-end 4S/8S server processor lines	Core count	Chipset	Proc. socket
Truland MP	Pentium 4 MP Prescott	90 nm	3/2005	90 nm Pentium 4 MP (Potomac)	1C	E8500 + ICH5	LGA 604
	Pentium 4 Presc.	90 nm	11/2005	7000 (Paxville MP)	2x1C	E8501 + ICH5	
	Pentium 4 Presc.	65 nm	8/2006	7100 (Tulsa)	2x1C		
Caneland	Core2	65 nm	9/2007	7200 (Tigerton DC) 7300 (Tigerton QC)	2C 2x2C	E7300 (Clarksboro) + 631x/632x ESB	LGA 604
	Penryn	45 nm	9/2008	7400 (Dunnington)	6C		
Boxboro-EX	Nehalem	45 nm	3/2010	7500 (Beckton/ Nehalem-EX)	8C	7500 + ICH10	LGA 1567
	Westmere	32 nm	4/2011	E7-8800 (Westmere-EX)	10C		
	Sandy Bidge	32 nm					
Brickland	Ivy Bridge	22 nm	2/2014	E7-8800 v2 (Ivy Bridge-EX)	15C	C602J (Patsburg J)	LGA 2011-1
	Haswell	22 nm	5/2015	E7-8800 v3 (Haswell-EX)	18C		
	Broadwell	14 nm	6/2016	E7-8800 v4 (Broadwell-EX)	24C		
Purley	Skylake	14 nm	7/2017	Platinum 8100 (Skylake-SP)	28C	C620 (Lewisburg)	LGA 3647

a) Evolution of core counts in Intel's high-end 4S/8S multicore server processors -2



a1) Evolution of core counts in Intel's high-end 4S multicore server processors before L3 caches emerged -1



*

a1) Evolution of core counts in Intel's high-end 4S multicore server processors before L3 caches emerged -2

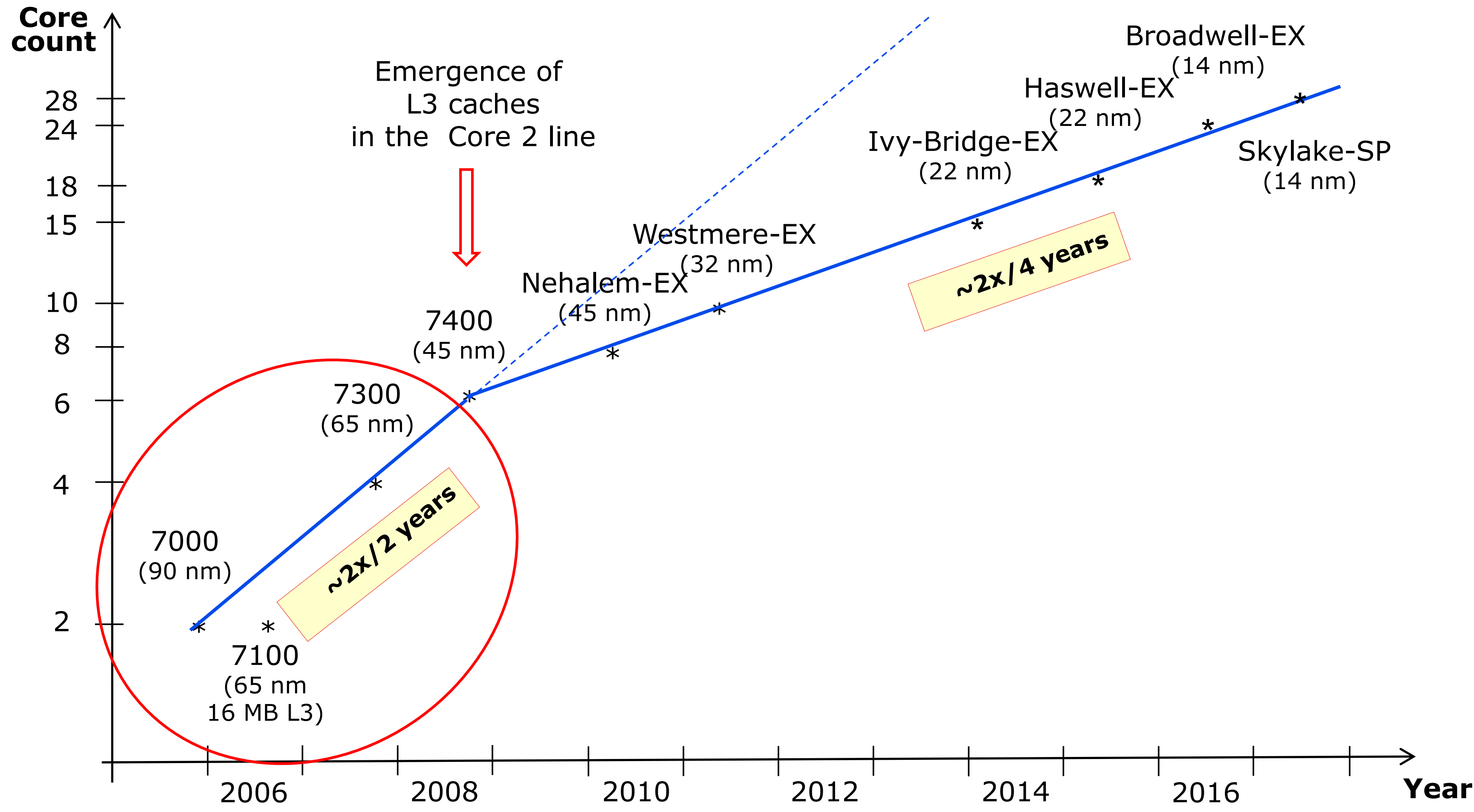
- As the above Figure shows core count roughly doubled every two years in early, L3 caache-less server processors.
- In light of Moore's rule this is obvious, since IC technology developments allowed to double transistor counts and thus also core counts roughly every two years as long as L3 caches did not occupy large parts of the silicon area.

Remark

In the Figure indicating the grows of core counts, the 7100 line don't fit to the line describing the raise of core counts in early L3 cache-less high-end servers.

*

a1) Evolution of core counts in Intel's high-end 4S multicore server processors before L3 caches emerged -3



The reason for this discrepancy is straightforward, since the Pentium 4-based 7100, albeit manufactured on a smaller feature size (65 nm) included already a large, 16 MB L3 cache that reduced the silicon area available for the cores and thus limited the core count.

*

a2) Evolution of core counts in Intel's high-end 4S multicore server processors after L3 caches emerged -1

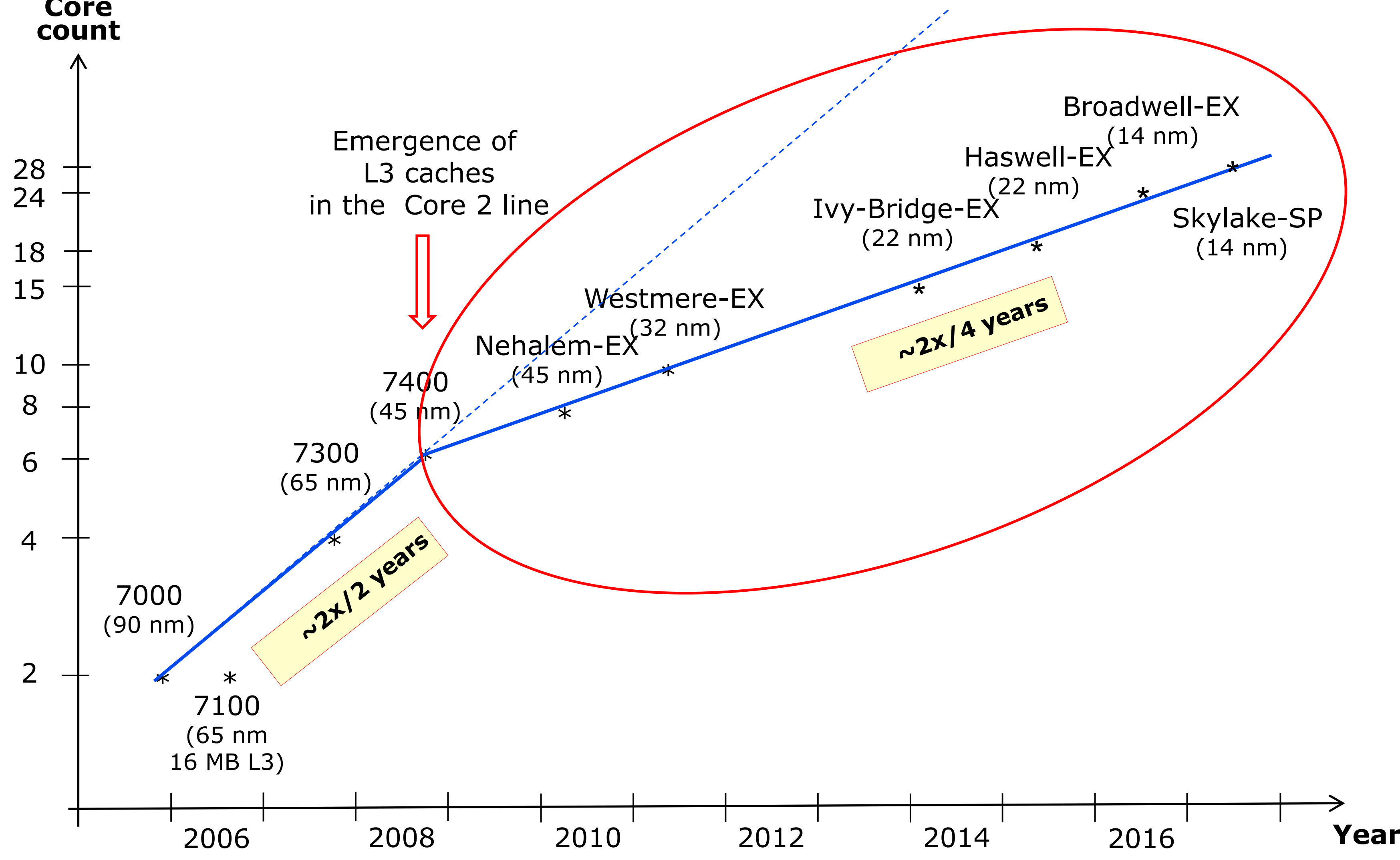


Figure: Core counts in Intel's high-end multicore 4S/8S server processors

*

Introduction of L3 caches in Intel's high-end 4S/8S multicore server processor lines

	Core	Techn.	Intro.	High-end 4S/8S server processors	Core count	L3 MB	Chipset	Proc. socket
Truland MP	Pentium 4 MP Prescott	90 nm	3/2005	90 nm Pentium 4 MP (Potomac)	1C	8	E8500 + ICH5	LGA 604
		90 nm	11/2005	7000 (Paxville MP)	2x1C	--	E8501 + ICH5	
		65 nm	8/2006	7100 (Tulsa)	2x1C	16		
Caneland	Core2	65 nm	9/2007	7200 (Tigerton DC) 7300 (Tigerton QC)	2C 2x2C	--	E7300 (Clarksboro) + 631x/632x ESB	LGA 604
	Penryn	45 nm	9/2008	7400 (Dunnington)	6C	16		
Boxboro-EX	Nehalem	45 nm	3/2010	7500 (Beckton/ Nehalem-EX)	8C	8x3	7500 + ICH10	LGA 1567
	Westmere	32 nm	4/2011	E7-8800 (Westmere-EX)	10C	10x3		
Brickland	Sandy Bidge	32 nm					C602J (Patsburg J)	LGA 2011-1
	Ivy Bridge	22 nm	2/2014	E7-8800 v2 (Ivy Bridge-EX)	15C	15x2.5		
	Haswell	22 nm	5/2015	E7-8800 v3 (Haswell-EX)	18C	18x2.5		
	Broadwell	14 nm	6/2016	E7-8800 v4 (Broadwell-EX)	24C	24x2.5		
Purley	Skylake	14 nm	7/2017	Platinum 8100 (Skylake-SP)	28C	28x2.5	C620 (Lewisburg)	LGA 3647

Introducing an L3 cache in the Penryn-based Xeon 7400 MP (2008)

As the Figure shows the 16 MB L3 (LLC) covers almost the same die area as the 6 cores.

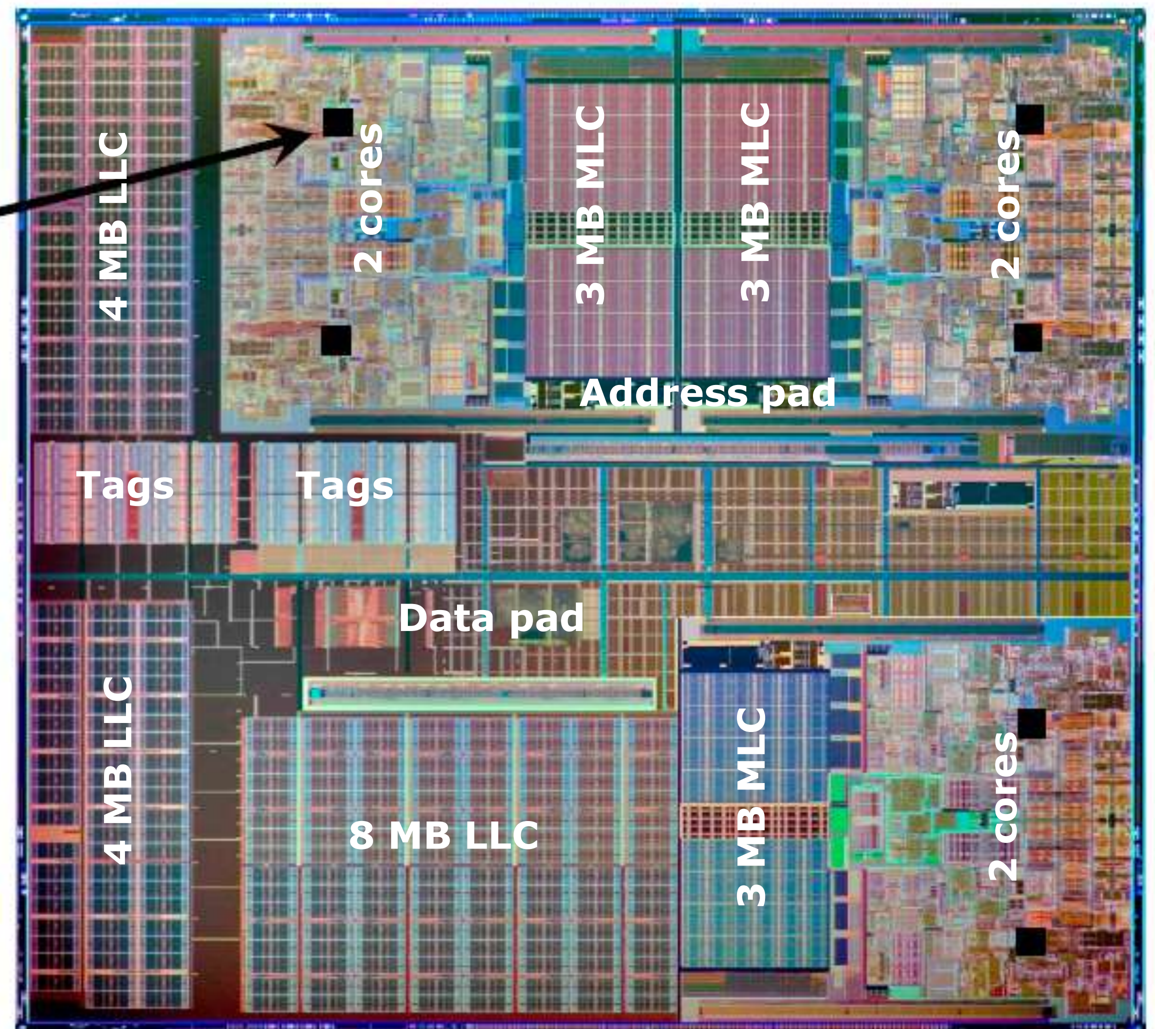


Figure Die plot of the Penryn-based Xeon 7400 MP [96]

a2) Evolution of core counts in Intel's high-end 4S/8S multicore server processors after L3 caches emerged -2

The Figure showing the evolution of core counts indicates that after L3 caches appeared core counts raised in a slower rate as before, approximately by doubling every four years.

This rate results from the similar rate of raising transistor counts, as the subsequent Table and the related Figure show.

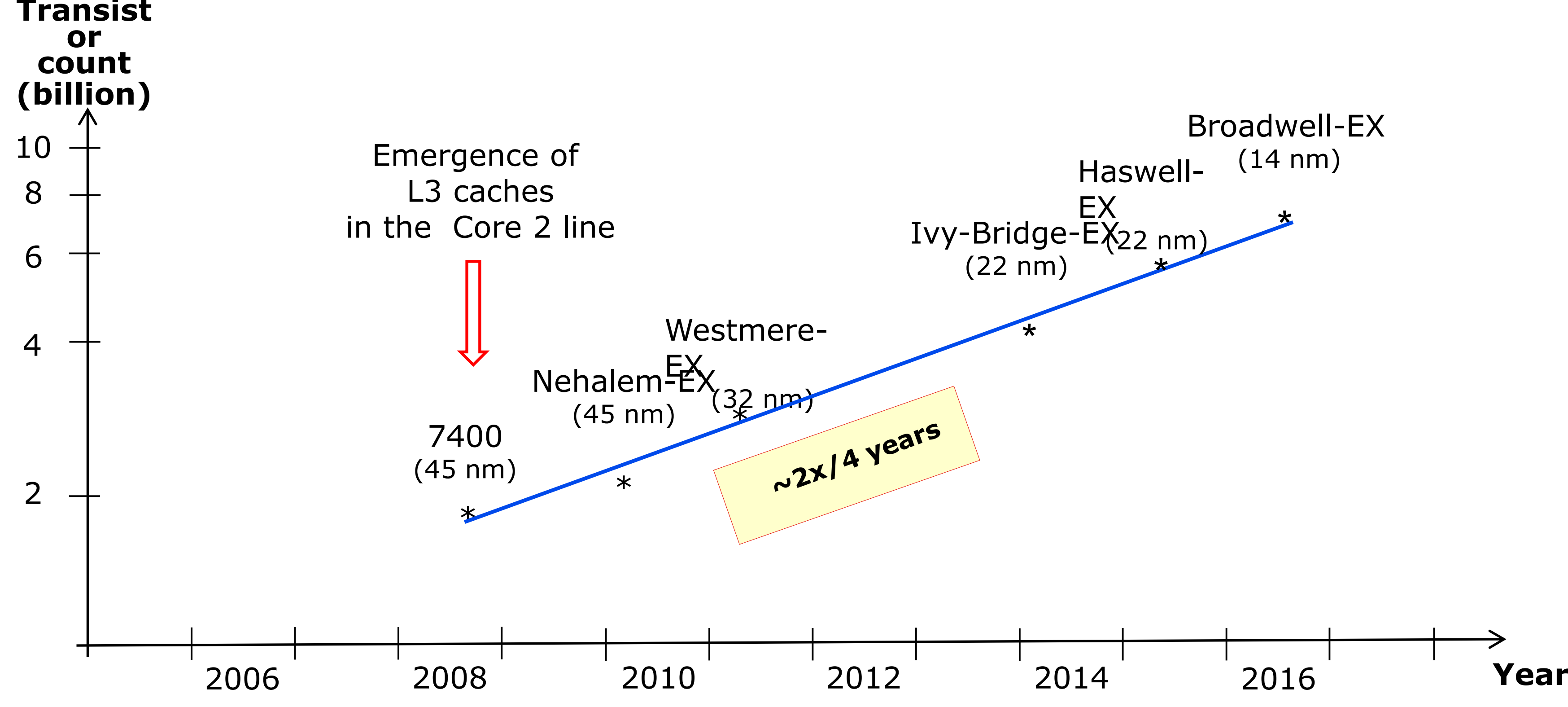
Family	Technology	Core count	L3 cache	Transistors (million)	Die area (mm^2)
7300	65 nm	2x2	no L3	582	286
7400	45 nm	6	16 MB	1900	503
Nehalem-EX	45 nm	8	8x3 MB	2300	513
Westmere-EX	32 nm	10	10x3 MB	2600	512
Ivy Bridge-EX	22 nm	15	14x2.5 MB	4310	541
Haswell-EX	22 nm	18	18x2.5 MB	5700	662
Broadwell-EX	14 nm	24	24x2.5 MB	7200	456

Table: Specific data of Intel's Core 2 based server processor dies

(Source: Intel's data)

*

Evolution of transistor counts of Intel's Core 2 based server processor dies



The slower rate of raising transitor counts can presumable be attributed to the limited power budget for each technology node.

*

Remark -1

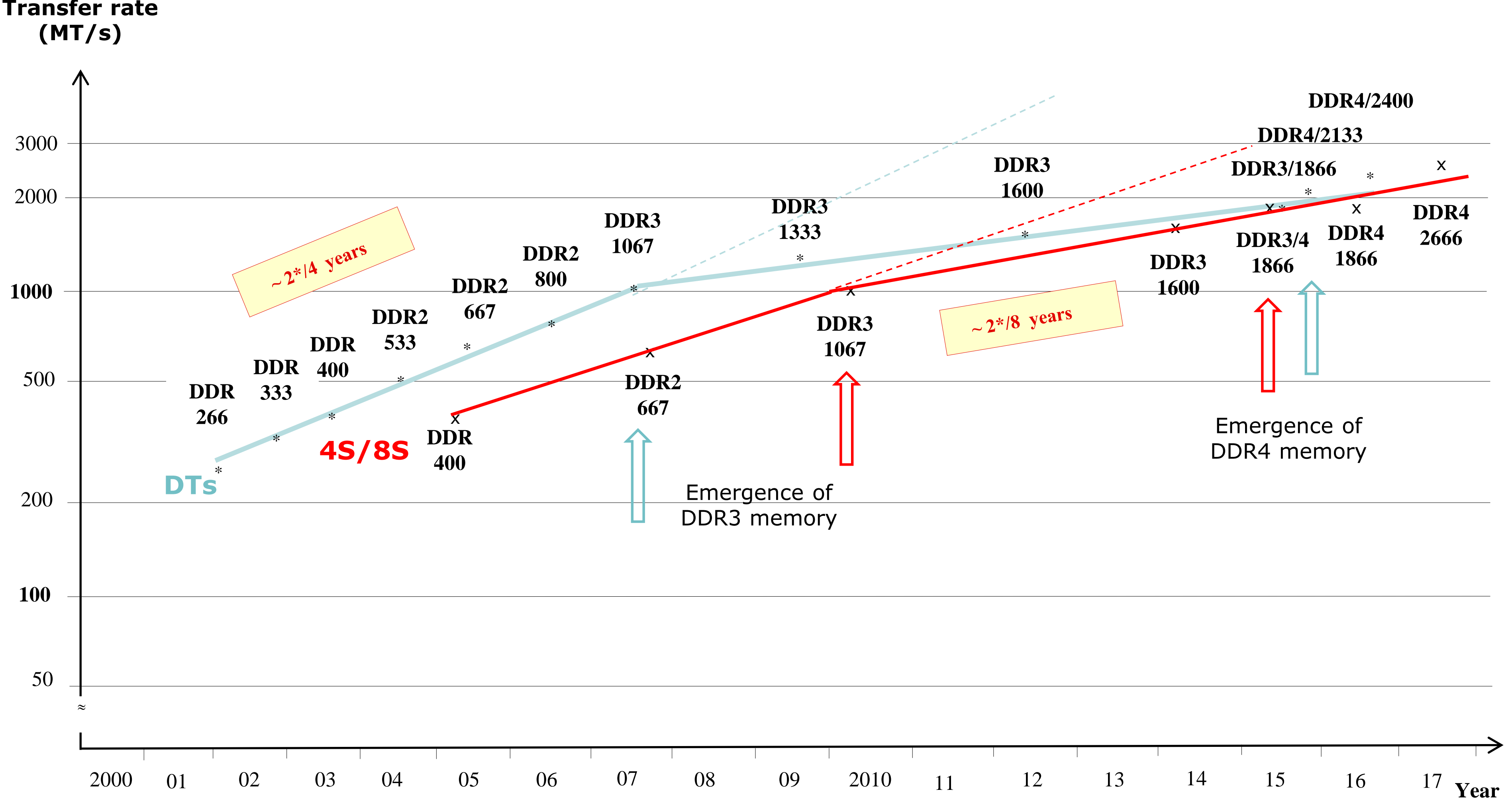
- There is an anomaly in respect to the 6-core Xeon 7400 since it lies on the evolution path of the L3-less processors albeit it includes already a 16 MB large L3 cache.
- The reason for this is simple, as this processor has been fabricated by 45 nm technology vs. 65 nm on an almost doubled silicon area (503 mm^2 vs. 286 mm^2) compared to the foregoing Xeon 7300, allowing nearly to qudruple the transistor count from the previous 582 million to 1.9 billion.

  In this way the large available transistor count did not restrict the core count.
- Here we note that subsequent processors were fabricated on about the same silicon area as the Xeon 7400 thus these processors (as seen in the Table below), and therefore followed the evolution path with the reduced slope.

Family	Technology	Core count	L3 cache	Transistors (million)	Die area (mm^2)
7300	65 nm	2x2	no L3	582	286
7400	45 nm	6	16 MB	1900	503
Nehale-EX	45 nm	8	8x3 MB	2300	513
Westmere-EX	32 nm	10	10x3 MB	2600	512
Ivy Bridge-EX	22 nm	15	14x2.5 MB	4310	541
Haswell-EX	22 nm	18	18x2.5 MB	5700	662
Broadwell-EX	14 nm	24	24x2.5 MB	7200	456

Table: Specific data of Intel's Core 2 based server processor dies

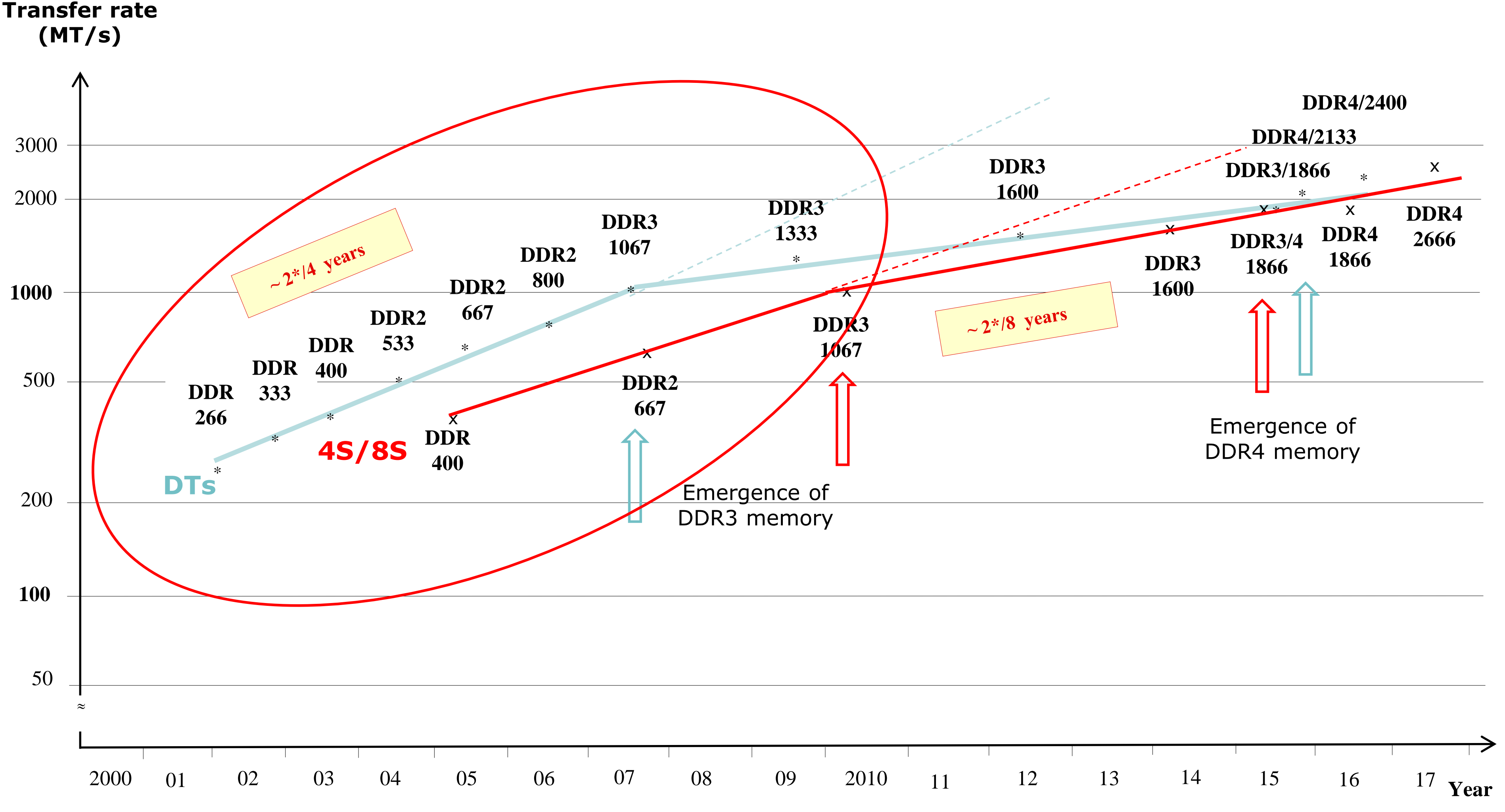
b) Evolution of memory transfer rates in Intel's DTs and HE multicore 4S/8S servers



HE: High-End

*

Evolution of DDR/DDR2 memory transfer rates in Intel's DTs and HE 4S/8S servers -1



HE: High-End

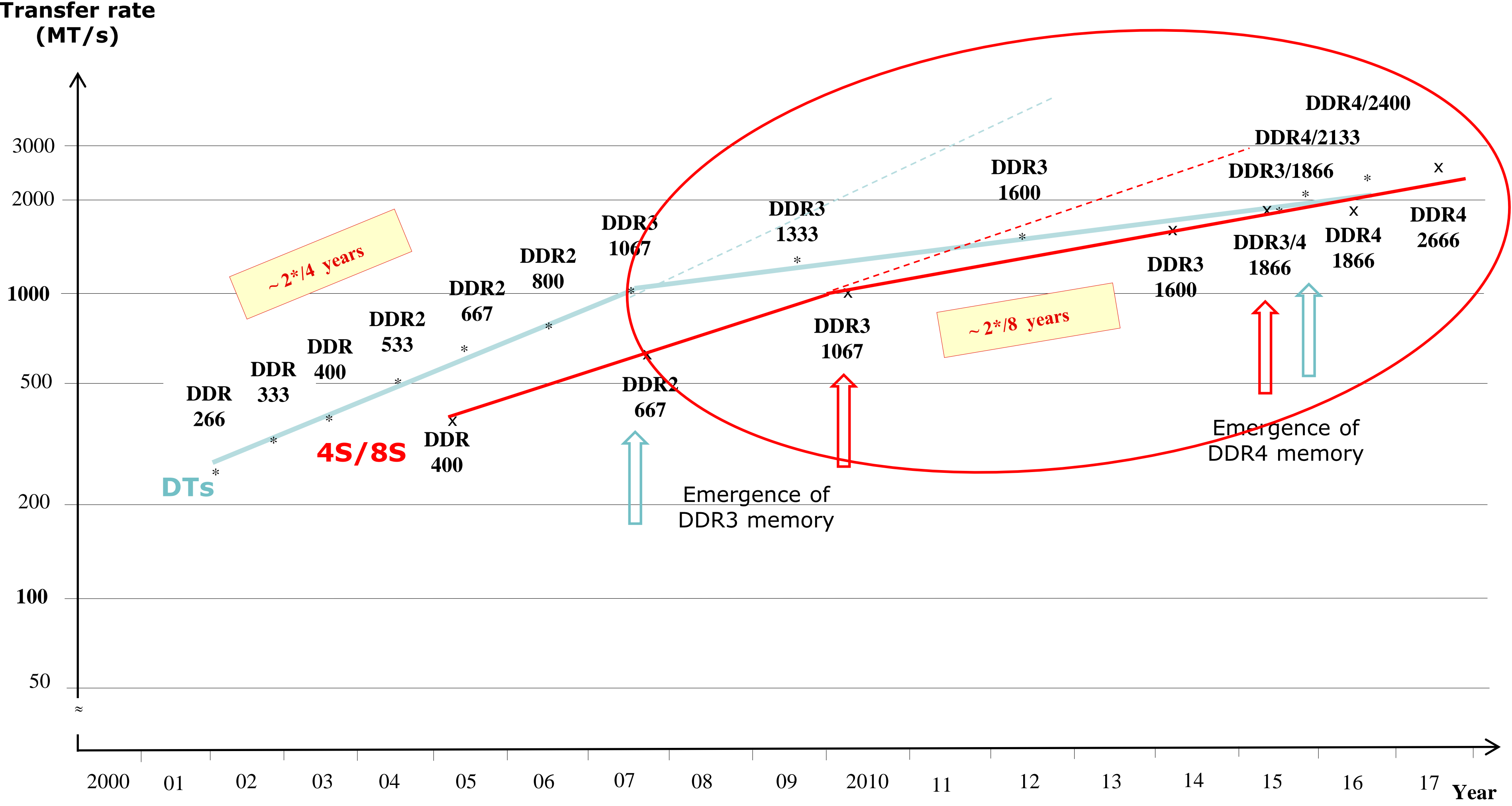
*

Evolution of DDR/DDR2 memory transfer rates in Intel's DTs and HE 4S/8S servers -2

- As the Figure shows, in the course of evolution transfer rates of DDR2 memories doubled roughly every four years, that is roughly half as fast as transfer rates of core counts of high-end 4S/8S server processors.

*

Evolution of DDR3/DDR4 memory transfer rates in Intel's DTs and HE 4S/8S servers -1



HE: High-End

*

Evolution of DDR3/DDR4 memory transfer rates in Intel's DTs and HE 4S/8S servers -2

As the above Figure shows, after DDR3 memory emerged rate of memory transfer rates slowed down to doubling rougly every eight years due to the higher complexity of the technology.

*

2.2 The memory bandwidth bottleneck

2.2 The memory bandwidth bottleneck

a) The need for preserving the per core memory bandwidth in multicore processors

- In servers every core may run independently from each other, consequently, the memory bandwidth needs to be linearly scaled with the core count, in other words

  while processor generations evolve the per core memory bandwidth should be preserved.

- The per core memory bandwidth of a socket (BW_M/core) amounts to

$$BW_M/\text{core} = n_M * w * T_M / n_C$$

with n_M: Number of memory channels
w: Width of a memory channel (8 byte)
T_M: Transfer rate of the memory (e.g 2.4 GT/s)
n_C: Core count

*

2.2 The memory bandwidth bottleneck (2)

What happens when over server generations the memory transfer rate raises slower than the core count?

If in subsequent processor generations the memory transfer rate (T_M) rises slower than the core count (n_C), then in the expression of the per core memory bandwidth (BW_M/core

$$BW_M/core = n_M * w * T_M / n_C$$

the ratio of T_M / n_C decreases and consequently, the per core memory bandwidth decreases and a memory bandwidth gap arises.

*

How can the per-core memory bandwidth be preserved when over generations the memory transfer rate raises slower than the core count?

When the ratio of T_M / n_C decreases the per-core memory bandwidth (BW_M/core) may be preserved if n_M becomes suitable increased, as the expression for BW_M/core reveals:

$$BW_M/\text{core} = n_M * w * T_M / n_C$$

In other words

if in a processor family the per core memory bandwidth should be preserved over generations the emerging memory bandwidth gap needs to be compensated by appropriately raising the number of memory channels (n_M) on the platform.

Here we assume that the width of the memory channels (w) remains unchanged.

*

The number of memory channels (n_M) needed per socket for preserving the per core memory bandwidth of an n-core high-end 4S/8S server

- Assuming the discussed core count and memory transfer growth rates it has been shown [] that a high-end server with n_C cores needs approximately $n_M = \sqrt{2} * \sqrt{n}$ memory channels per socket if the per core memory bandwidth should be preserved related to typical early single core 4S server implementations.
- The required memory channel numbers per socket are shown in the Table below.

n_C	n_M
2	2.0
4	2.8
6	3.5
8	4.0
10	4.5
12	4.9
14	5.3
16	5.6
20	6.3
24	6.9
32	8.0
48	9.8
64	11.3

Table: The number of memory channels (n_M) needed per socket for preserving the per core memory bandwidth of a high-end server with n_C cores

*

2.2 The memory bandwidth bottleneck (5)

Note

Providing enough memory channels and thus memory bandwidth per socket is one of the key challenges in designing server processors and platforms since multicores emerged and core counts continuously raise due to electrical and power constraints, as discussed in the subsequent Sections.

Example: Raising the number of memory channels (n_M) in Intel's high-end 4S/8S servers

In order to preserve the per core memory bandwidth the number of memory channels rose rapidly from 1 per socket to 8 per socket in Intel's high-end server platforms, as shown in the next Table.

*

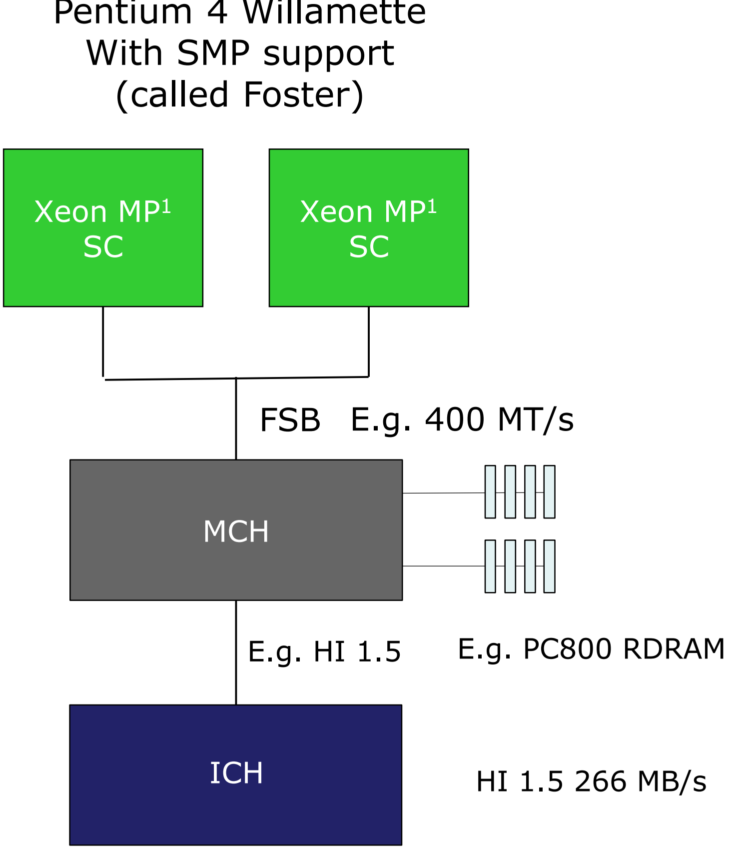
No. and max. transfer rate of memory channels in Intel's high-end server lines

Platform	Platform topology	Date of intro.	Processor	Technology	Core count (up to)	No. and max. speed of memory channels/socket	Kind of connecting memory Socket
Truland MP	SMP w/dual FSBs	3/2005	90 nm Pentium 4 MP (Potomac)	90 nm	1C	1x DDR2-400/socket	4 serial channels (IMI) via MCH LGA 604
		11/2005	Xeon 7000 (P4) (Paxville MP)	90 nm	2x1C		
		8/2006	Xeon 7100 (P4) (Tulsa)	65 nm	2C		
Caneland	SMP w/Quad FSBs	9/2007	Xeon 7200 (Core 2) Tigerton DC	65 nm	2C	1x DDR2-667/socket (FB-DIMM)	4 serial FB-DIMM channels via MCH LGA 604
		9/2007	Xeon 7300 (Core 2) Tigerton QC	65 nm	2x2C		
		9/2008	Xeon 7400 (Penryn) (Dunnington)	45 nm	6C		
Boxboro-EX	NUMA, fully connected for 4S platforms, partially connected for 8S platforms	3/2010	Nehalem-EX (Xeon 7500/ (Beckton)	45 nm	8C	8x DDR3-1067/socket	4 serial channels (SMI) /socket LGA 1567
		4/2011	Westmere-EX (E7-8800)	32 nm	10C		
Brickland		2/2014	Ivy Bridge-EX (E7-8800 v2)	22 nm	15C	8x DDR3-1600/socket	4 parallel channels (SMI2)/socket LGA-2011-1
		5/2015	Haswell-EX (E7-8800 v3)	22 nm	18C	8x DDR4-1866/socket	
		6/2016	Broadwell-EX (E7-8800 v4)	14 nm	24C	8x DDR4-1866/socket	
Purley		7/2017	Skylake-SP (Platinum 8100)	14 nm	28C	6x DDR4-2666/socket	6 direct attached DDR4 channels/socket LGA-3647

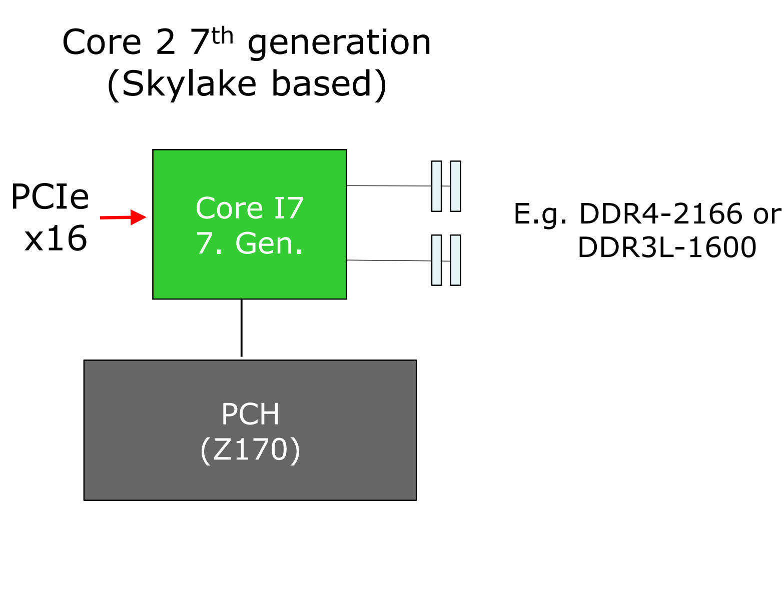
2.3 Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor

2.3 Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor -1

In early server and client platforms memory channels are typically directly connected to the MCH along with recent client platforms that usually attach memory directly to the processor or to processors in server platforms, as the examples below illustrate it.



High-level block diagram of an early Pentium 4 based DP server



High-level block diagram of a recent Core 2 Skylake based DT

Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor -2

- For directly connected standard DIMM channels each memory channel requires a large number of copper traces since these interconnects are parallel, 8-byte wide links.
- The exact number of copper traces needed depends on the memory type used, for actual numbers see the next Figure.

2.3 Maximum number of standard DRAM channels (3)

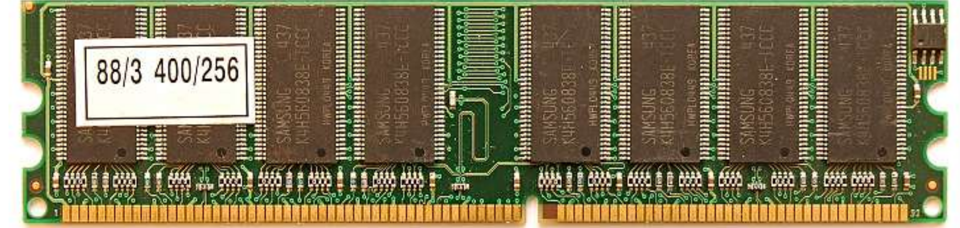
Pin counts of SDRAM to DDR4 DIMMs

All these DIMMs are 8-byte wide.

SDRAM (SDR)

168-pin

DDR

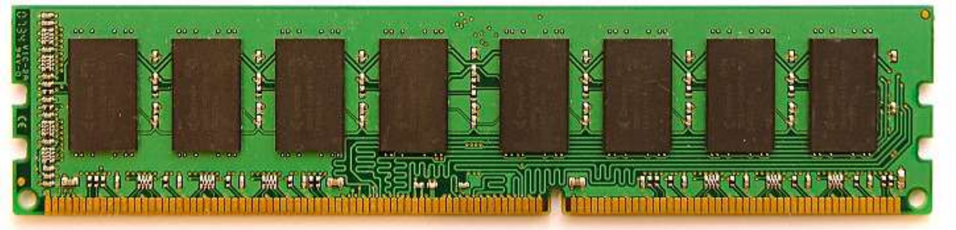


184-pin

DDR2

240- pin

DDR3



240-pin

DDR4

284-pin

Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor -3

The interconnects between the MCH or the processor socket and the DIMM socket(s) are implemented by dense copper trails on the mainboard, as seen below, typically equalized to reduce skews (differences in signal propagation times occurring on different long bit-lines).

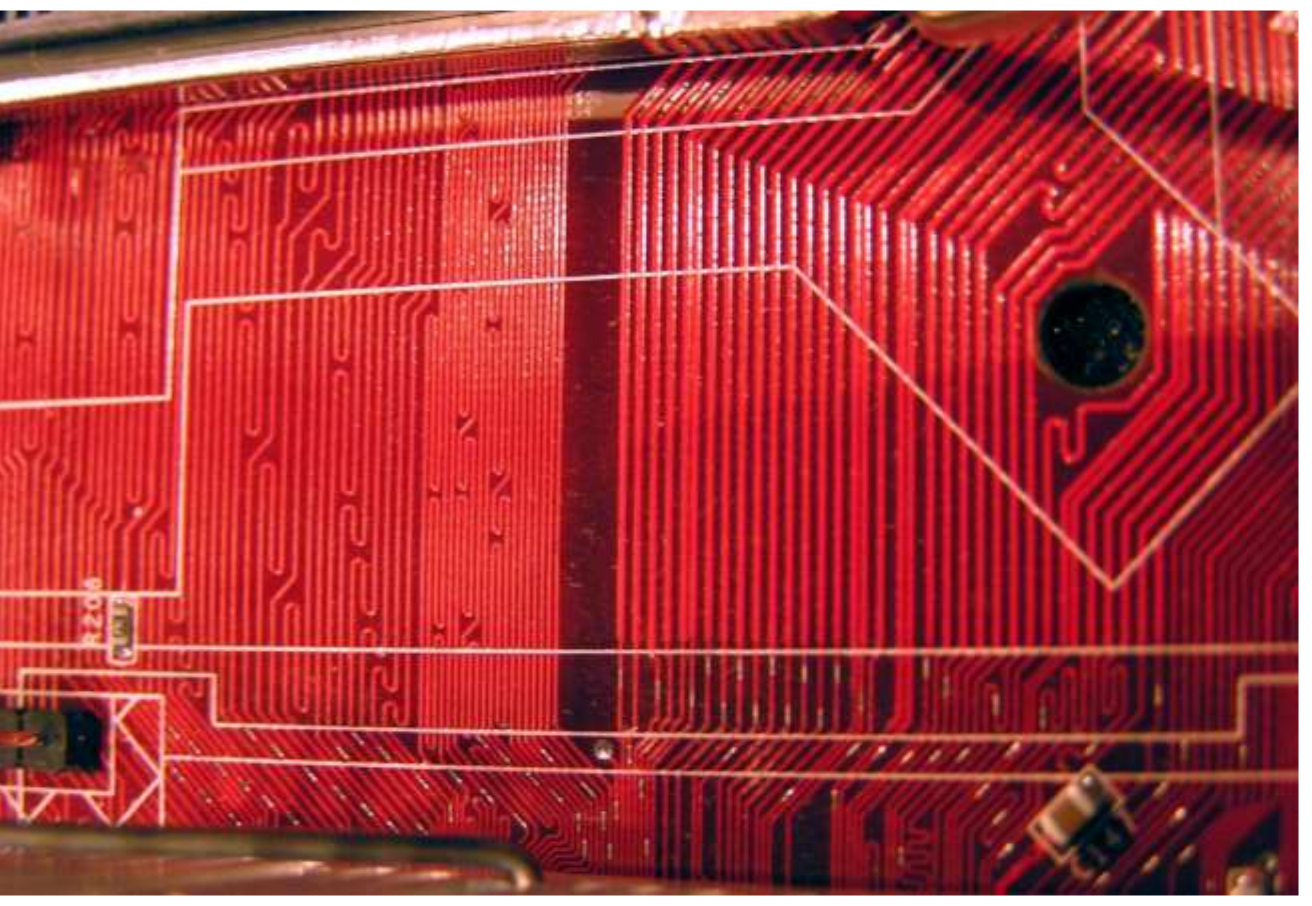


Figure: Copper trails interconnecting the MCH or processor socket and the DIMM socket(s)

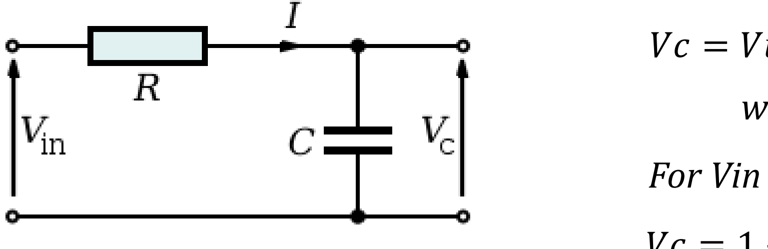
2.3 Maximum number of standard DRAM channels (5)

Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor -4

- For directly connected memory channels electrical constrains hardly limit the maximum number of attachable memory channels.
- This limitation arises basically out of the fact that the width and the distance of the copper trails from each other determine the electrical parameters and thus the temporal behavior of the interconnect, first of all its maximum transfer rate, as will be discussed subsequently.

Modelling the temporal behavior of denser copper traces interconnecting the MCH or the processor and the DIMMs

The temporal behavior of raising the voltage on DIMM traces can be modelled by a serial RC network with R representing the ohmic impedance and C the capacitance of a trace, as shown below.



$$Vc = Vin\ x\ (1 - e^{-t/\tau})$$

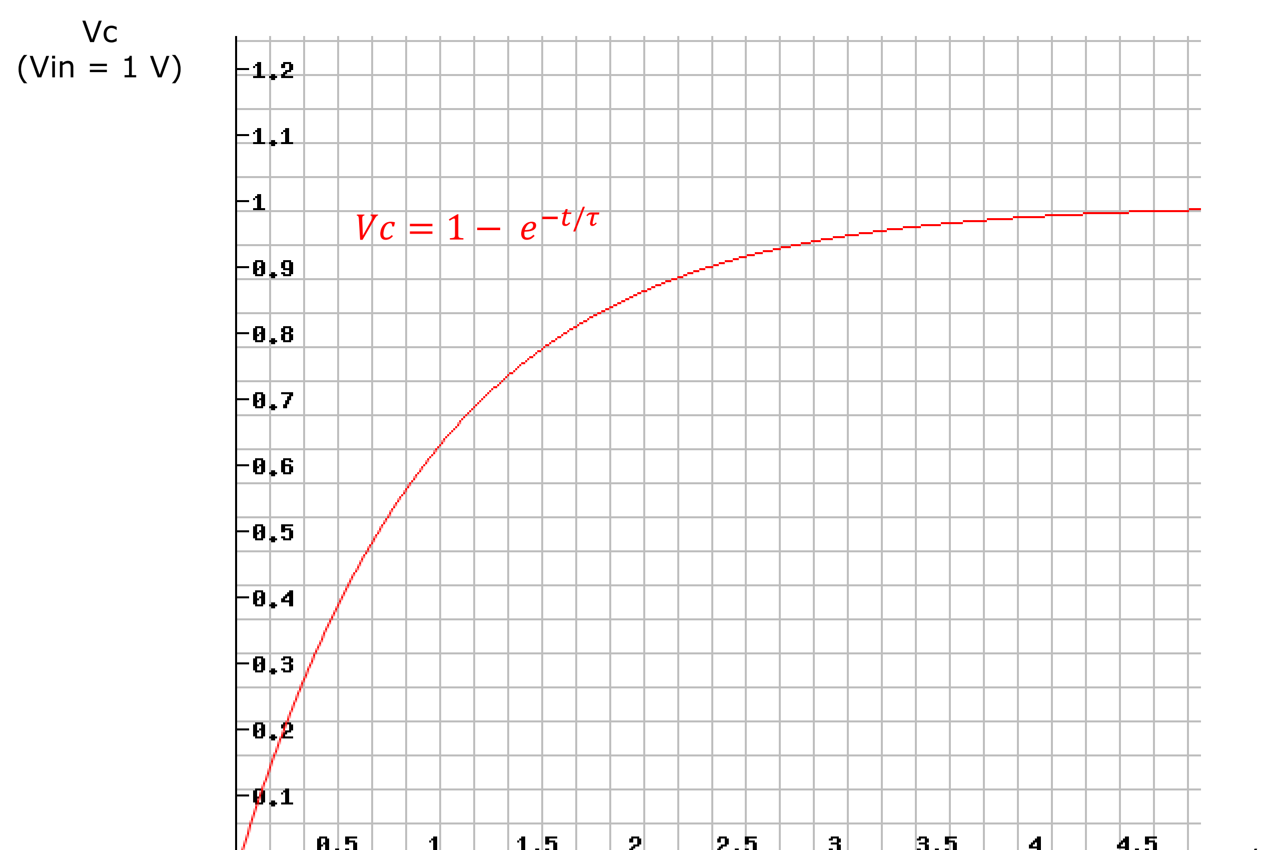
$$with\ \tau = R\ x\ C$$

For Vin = 1V:

$$Vc = 1 - e^{-t/\tau}$$

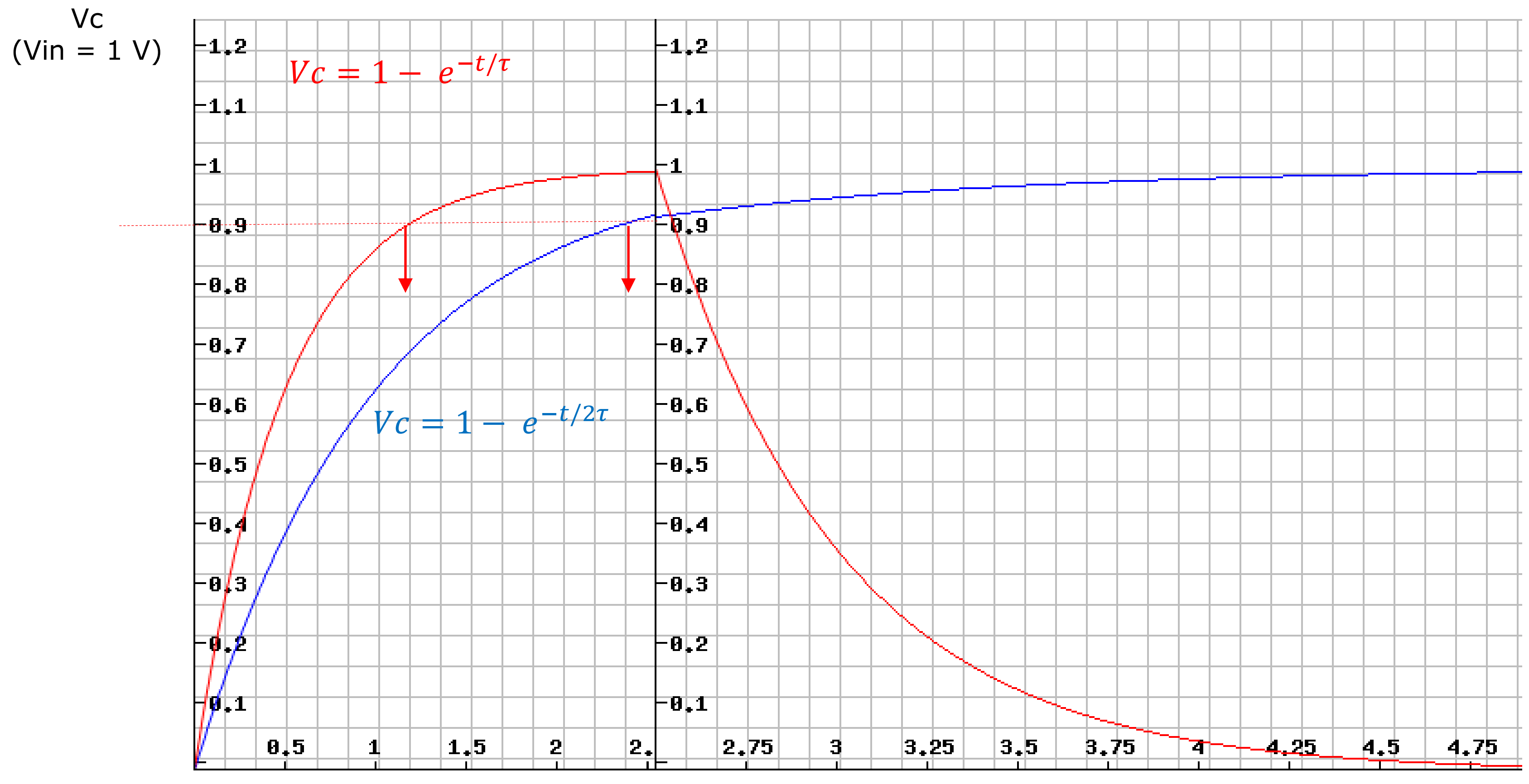
Figure: The circuit model and the related expression for Vc [135]

Raising the voltage on a capacitor of a serial RC network that models DIMM trails



τ = R * C

Raising the voltage on a capacitor of a serial RC circuit that models DIMM trails when τ is doubled



- Longer rise and fall times, however would reduce the feasible memory transfer rate.

Maximum number of standard DRAM channels that can directly be connected to the MCH or a processor -5

- As a consequence, if the given speed grade of the memory (e.g. DDR4-2400) should be fully exploited the maximum transfer rate should not limit it.
- In other words a given memory speed grade constrains the number of feasible copper trails and thus the number of memory channels that may directly be connected to an MCH or a processor.
- In this respect the physical dimensions of the MCH or processor socket are of paramount importance since larger sockets allow obviously, implementing more copper trails whereas smaller sockets less copper trails and thus more or less memory channels.
- We point out that for usual socket sizes of MCHs or client processors (of about 30 - 40 mm x 30 - 40 mm) typically up to two memory channels can directly be connected to an MCH or a processor to avoid memory speed reduction.
- Clearly, larger processor sockets will allow to connect more memory channels directly, as the next Figure indicates.

Number of directly attached memory channels and socket sizes of Intel's processors

P965 GMCH for Core 2 (2 cores)	**Westmere-EP** (up to 6 cores)	**Broadwell-EP** (up to 22 cores)	**Skylake-SP** (up to 28 cores)
34x34 mm	45x42.5 mm	58.5x51.0 mm	76x56 mm
BGA 1226	LGA 1366	LGA 2011-3	LGA 3647
2xDDR2 standard mem. channels on MCH	3xDDR3 standard mem. channels on processors	2x2 DDR4 standard mem. channels (2 on both side)	2x3 DDR4 standard mem. channels (3 on both side)

Note that physical/electrical constraints limit the number of attachable memory channels

Source of the pictures [144]

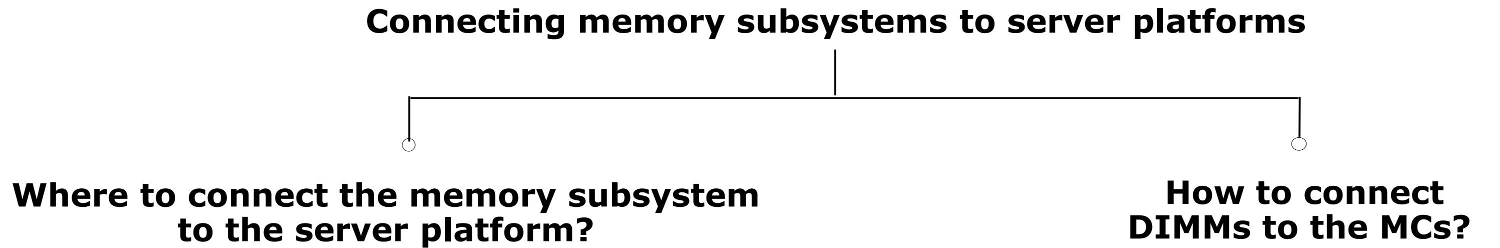
The driving force for the evolution of memory subsystems of high-end servers

- As pointed out in Section xx memory transfer rates rise in high-end 4S/8S servers slower than core counts and thus a memory bandwidth bottleneck arises unless appropriately more memory channels do not compensate it.
- This is however, not at all an easy task since the number of directly attached memory channels is restricted due to physical and electrical constraints, as discussed before e.g. more ore less to two if memory channels are connected via MCHs or client processors of typical sizes (of about 30 – 40 x 30 - 40 mm).
- The next Section is devoted entirely to this point by discussing the design space of connecting memory subsystems to server platforms in order to overview the possible approaches for raising the number of memory channels.

2.4 The basic design space of connecting memory subsystems to server platforms

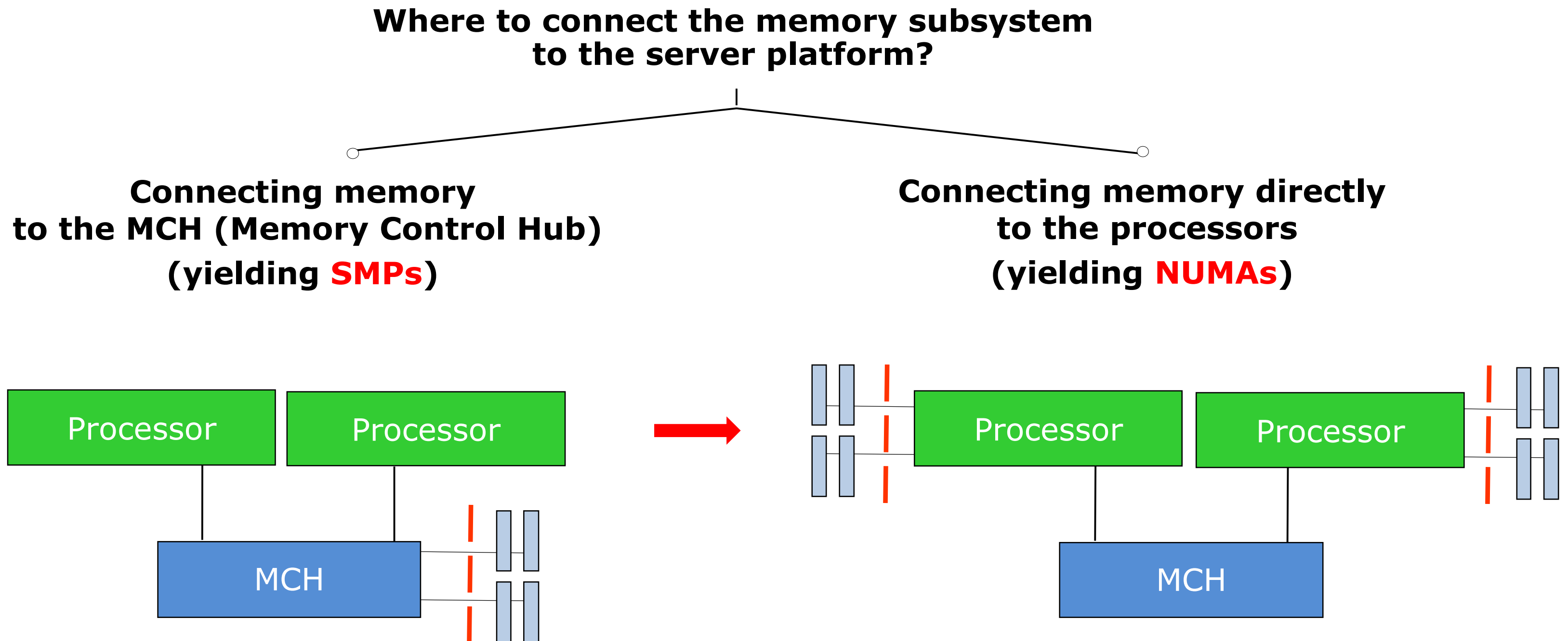
2.4 The basic design space of connecting memory subsystems to server platforms

It has two orthogonal dimensions, as shown below.



*

a) Where to connect the memory subsystems to the server platforms?



Memory is connected at a single point, i.e. centrally.

It causes severe bandwidth imitations
for multiprocessor server platforms.

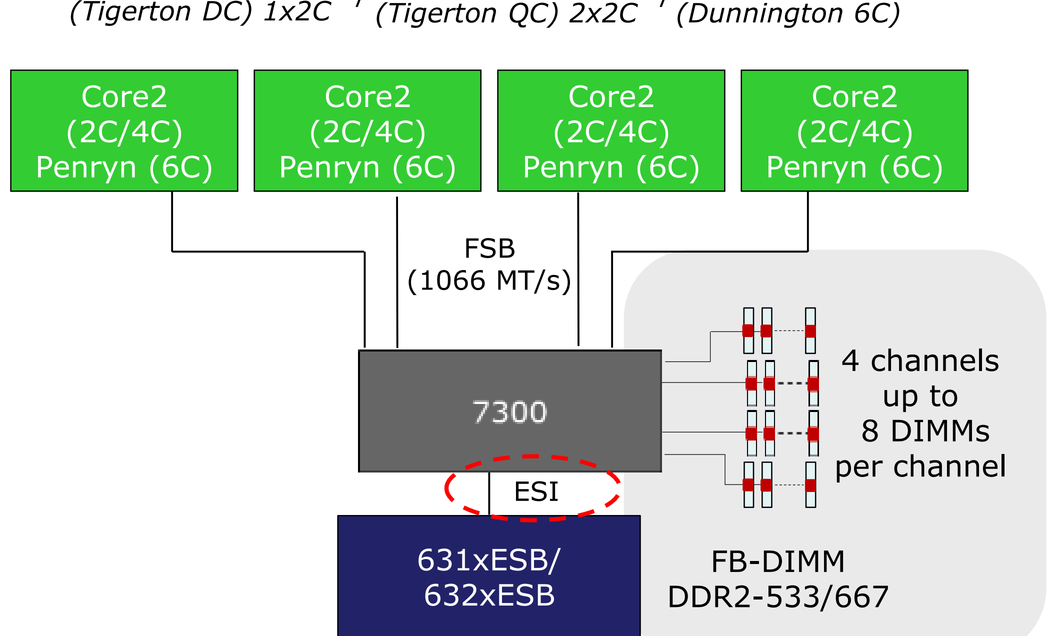
Memory is connected per processor, i.e. distributed.

It provides an inherent scaling
of the memory channels with the processor count,
i.e. it provides for higher processor counts
linearly higher memory bandwidth in
multiprocessor server platforms.

Figure: Options for the design aspect of "Where to connect the memory subsystem
to the server platform?"

*

Example 1: Connecting memory to the MCH in Intel's Core 2/Penryn-based high-end 4S server [138]

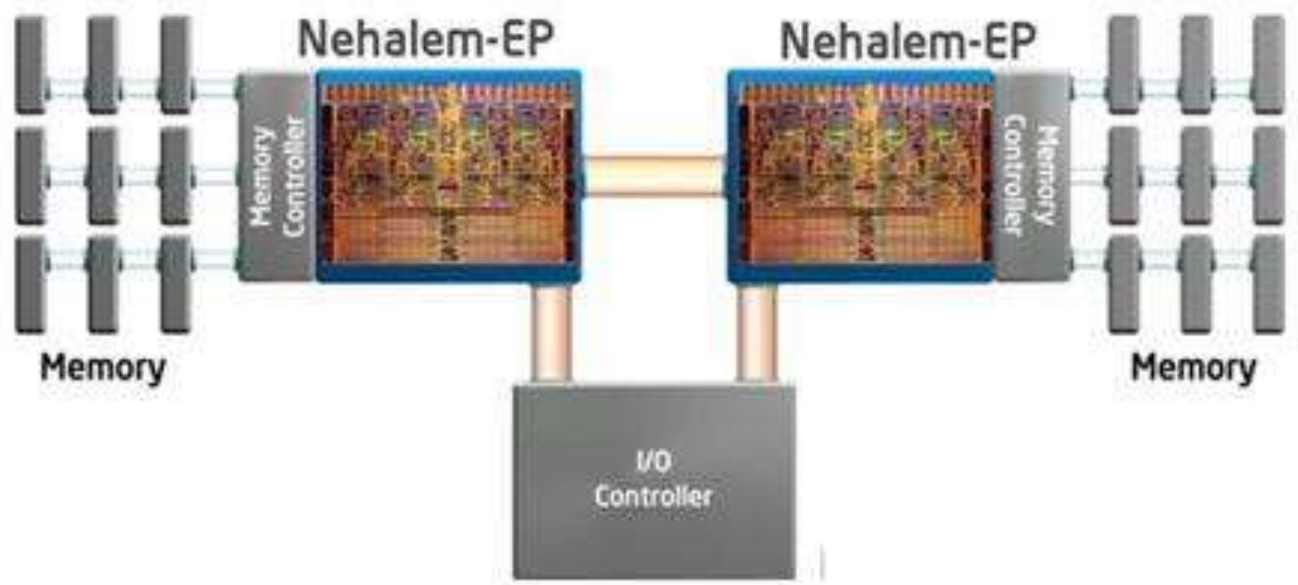


FB-DIMM:
Fully Buffered DIMM
(It will be discussed later in this Section)

Core 2/Penryn based
Caneland MP server platform (for up to 6 C)
(2007)

ESI: Enterprise System Interface
(4 PCIe lanes, 0.25 GB/s per lane (like the DMI interface, providing 1 GB/s transfer rate in each direction)

Example 2: Connecting memory directly to the processors in Intel's Nehalem-EP based DP server architecture [172]



Main features

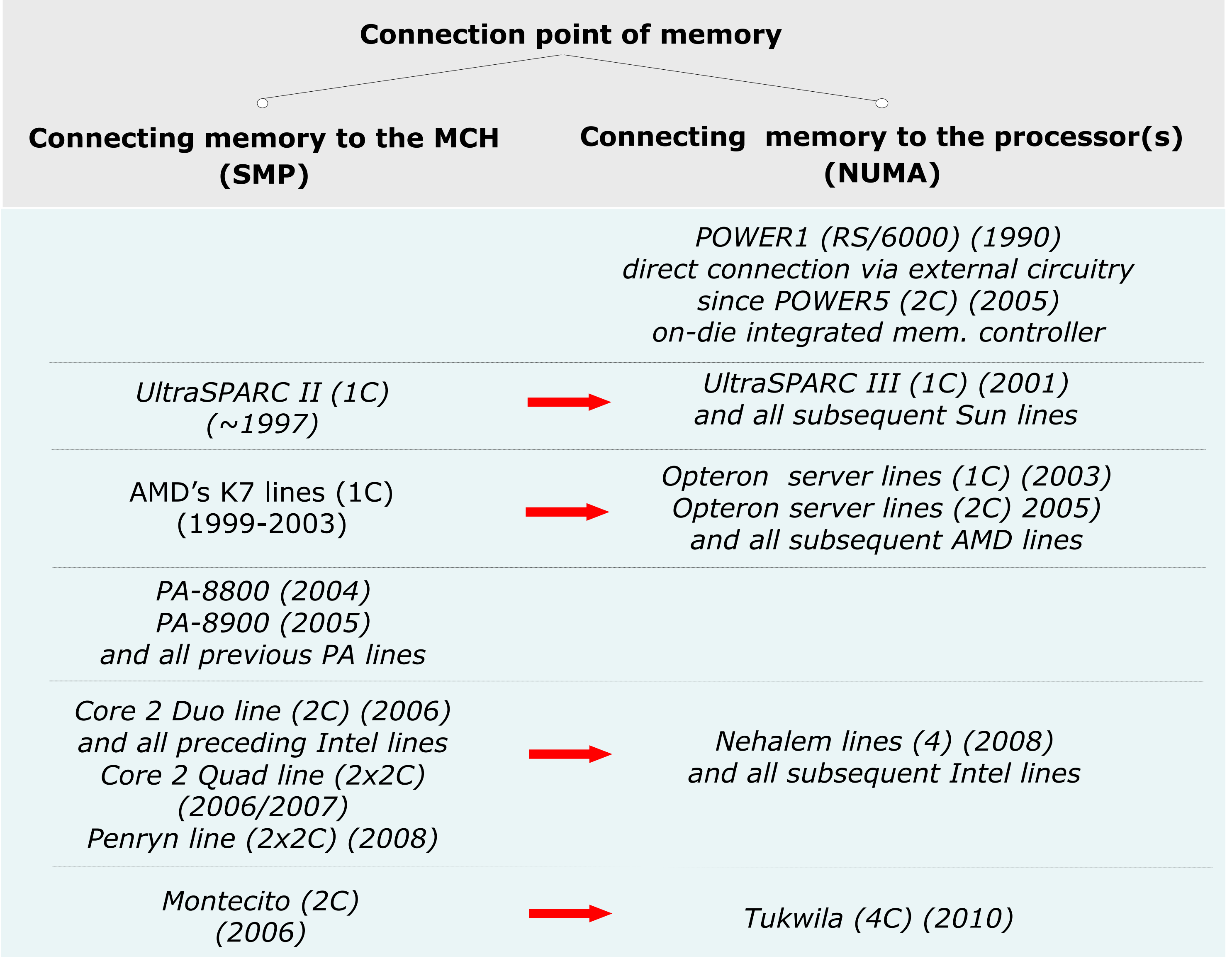
- 3 channels per socket
- Up to 3 DIMMs per channel (impl. dependent)
- DDR3-800, 1066, 1333,...

Nehalem-EP (Efficient Performance):
Designation of the server line

Remark

In order to connect 3 DDR3 memory channels to the processors the Nehalem-EP has larger sockets compared to the previous one (LGA1366 (42.5x45.0 mm) vs. the LGA771 (37.5x 37.5 mm))

Transition of the connection point of the memory in the course of processor evolution



b) How to connect DIMMs to the MCs?

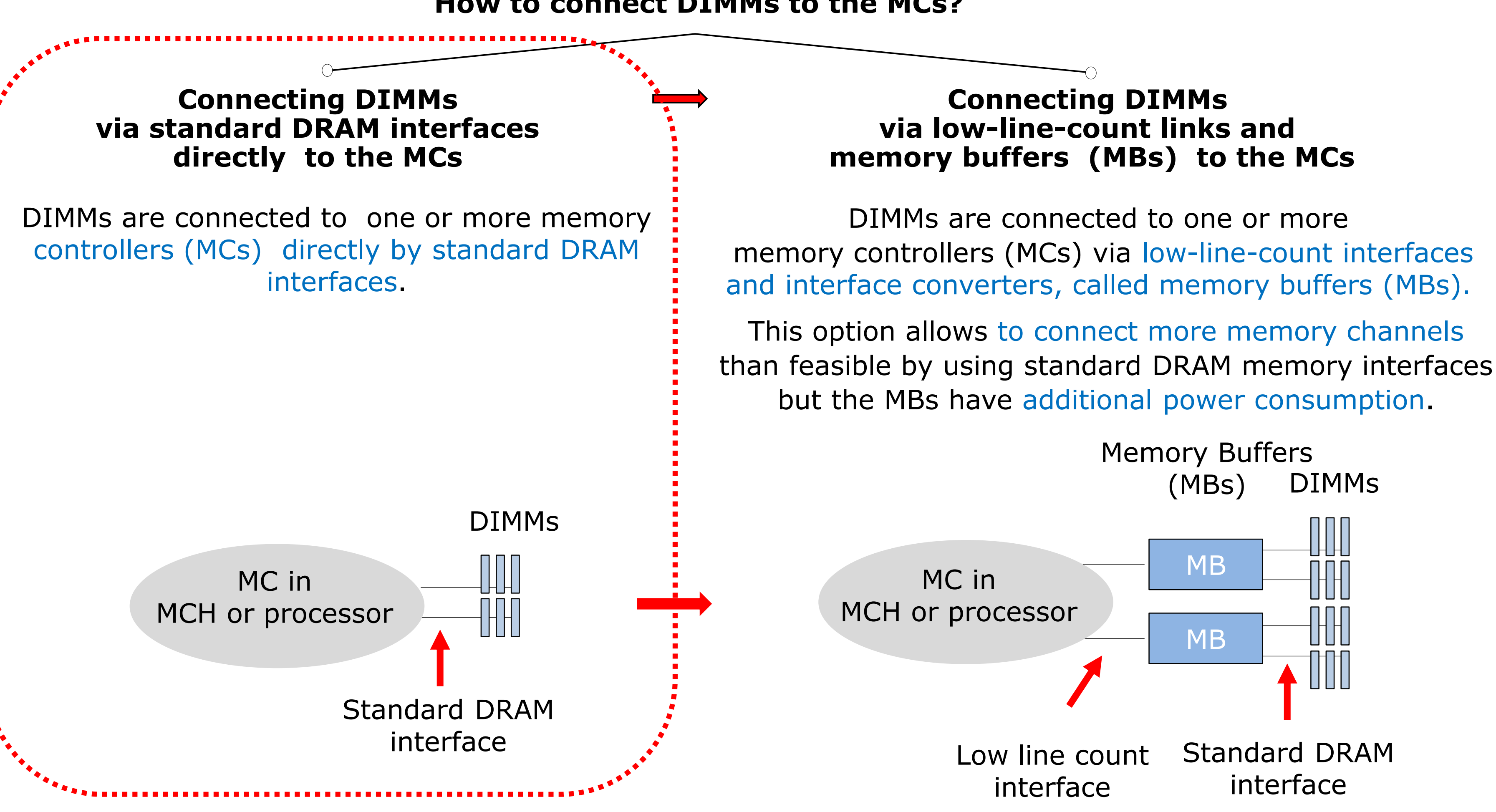
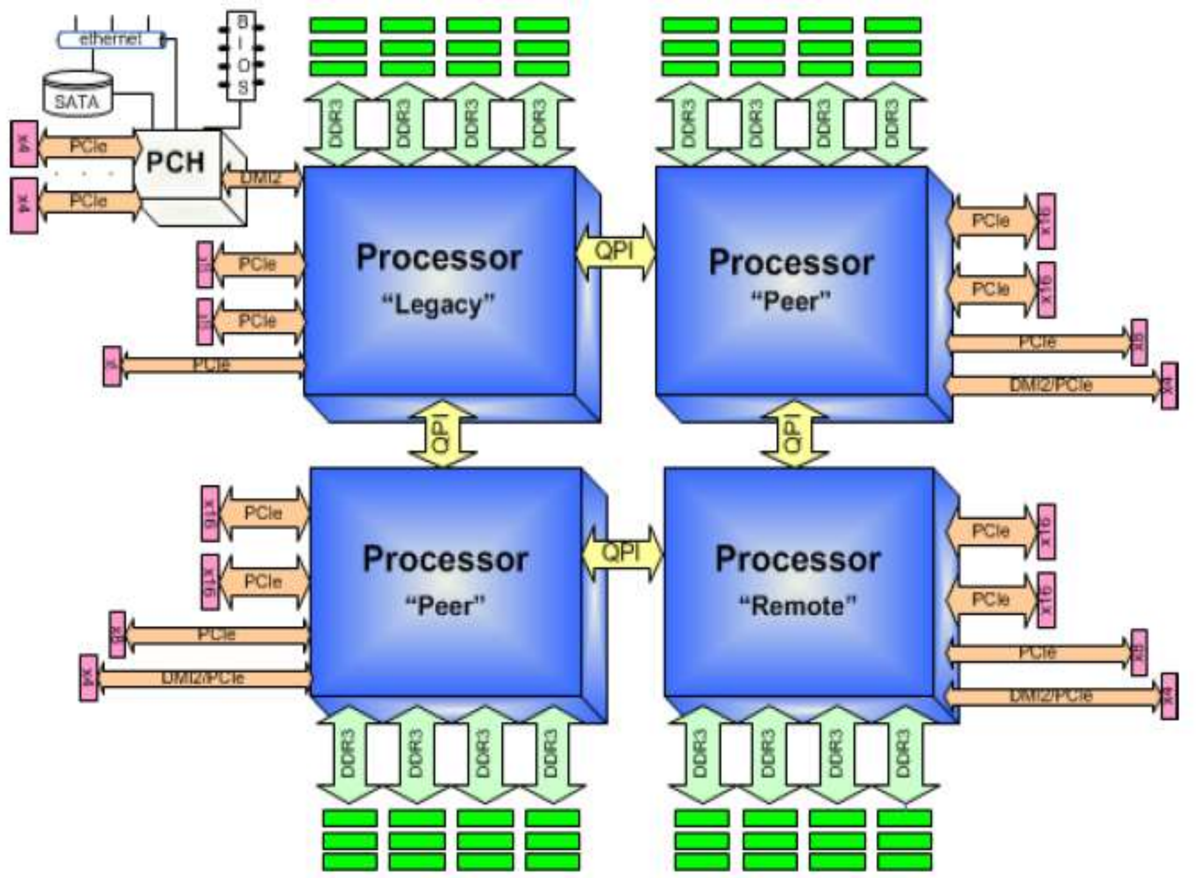


Figure: Options for implementing the connection path between the MCs and the DIMMs

*

2.4 The basic design space of connecting memory subsystems (7)

b1) Example: Connecting DIMMs via standard DRAM interfaces directly to the MCs in Intel's Sandy Bridge-EP based 4S server



58.5x51.0 mm

Figure: Sandy Bridge-EP based server [142]

Remark

In order to connect 4 DDR3 memory channels to the processors (each two on two opposite sides) the Sandy Bridge-EP has large sockets (LGA2011 measuring 58.5x51.0 mm).

b2) How to connect DIMMs to the MCs?

How to connect DIMMs to the MCs?

Connecting DIMMs via standard DRAM interfaces directly to the MCs

DIMMs are connected to one or more memory controllers (MCs) directly by standard DRAM interfaces.

Connecting DIMMs via low-line-count links and memory buffers (MBs) to the MCs

DIMMs are connected to one or more memory controllers (MCs) via low-line-count interfaces and interface converters, called memory buffers (MBs).

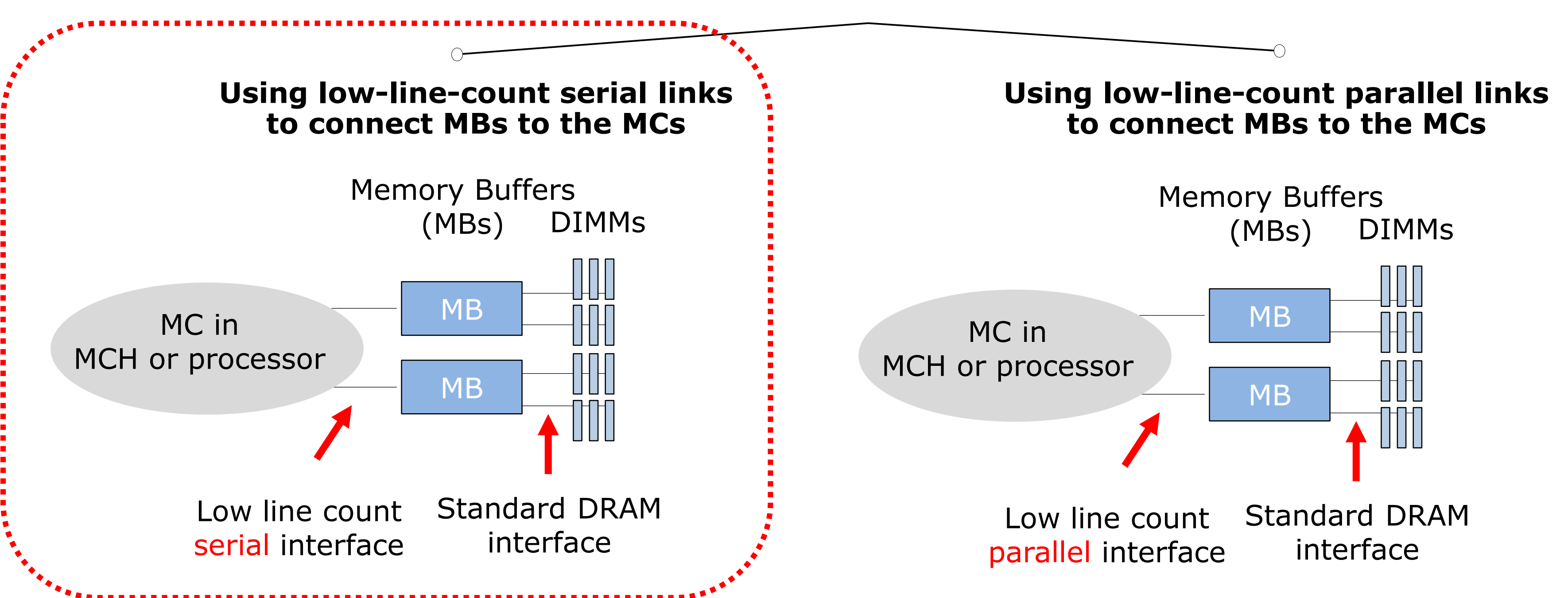
This option allows to connect more memory channels than feasible by using standard DRAM memory interfaces but the MBs have additional power consumption.

Figure: Options for implementing the connection path between the MCs and the DIMMs

*

b2) Connecting DIMMs via low line count links and memory buffers (MBs) to the MCs -1

There are two basic options for implementing low-line count links:



b2a) Using low-line-count serial links to connect MBs to the MCs -1

Connecting DIMMs via low-line-count links and memory buffers (MBs) to the MCs

Using low-line-count serial links to connect MBs to the MCs

The links are implemented as two unidirectional, Point-to-Point serial, packet based channels, typically with differential signaling.

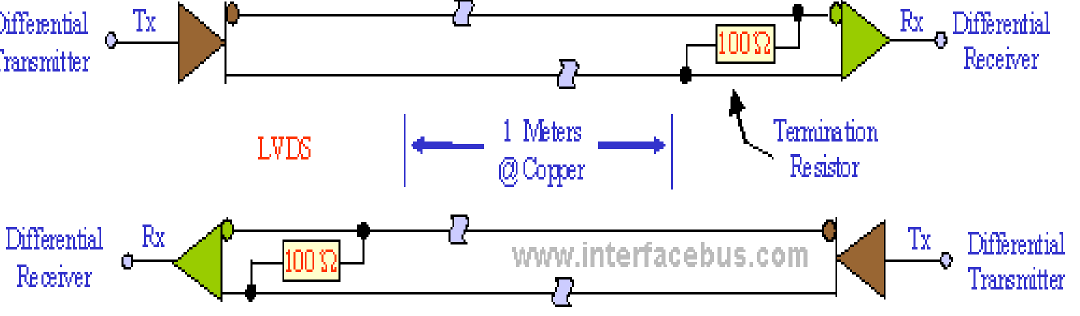


Figure: Example for differential signaling [170]

Drawback: The MBs have to perform serial/parallel conversion, this increases dissipation.

Used in Intel's early 4S or 8S Truland MP and Boxboro-EX platforms (2005 - 2011) for 2C to 10C

Implemented links (IMI, SMI) have about 70 lines (vs. 240 for DDR2/3).

Using low-line-count parallel links to connect MBs to the MCs

The links are implemented as 64-bit parallel, bidirectional channels typically with single-ended voltage reference signaling, termed VMSE (Voltage Mode Single Ended) signaling

Benefit: No need for serial/parallel conversion in the MBs, only interface conversion is required

Used in Intel's Brickland platform (2014 -2016) for 15C/18C/24C

Intel's implementation (the SMI 2 interface) comprises about 110 lines.

b2a) Using low-line-count serial links to connect MBs to the MCs -2

If using low-line count serial links there are two design choices for implementing MBs, as indicated below.

*

b2a1) Using on-board MBs (on the mainboard/riser card

If using low-line-count serial links how to implement MBs?

Using on-board MBs (on the mainboard/riser card)

The MBs are placed on the mainboard or riser cards

Example 1:

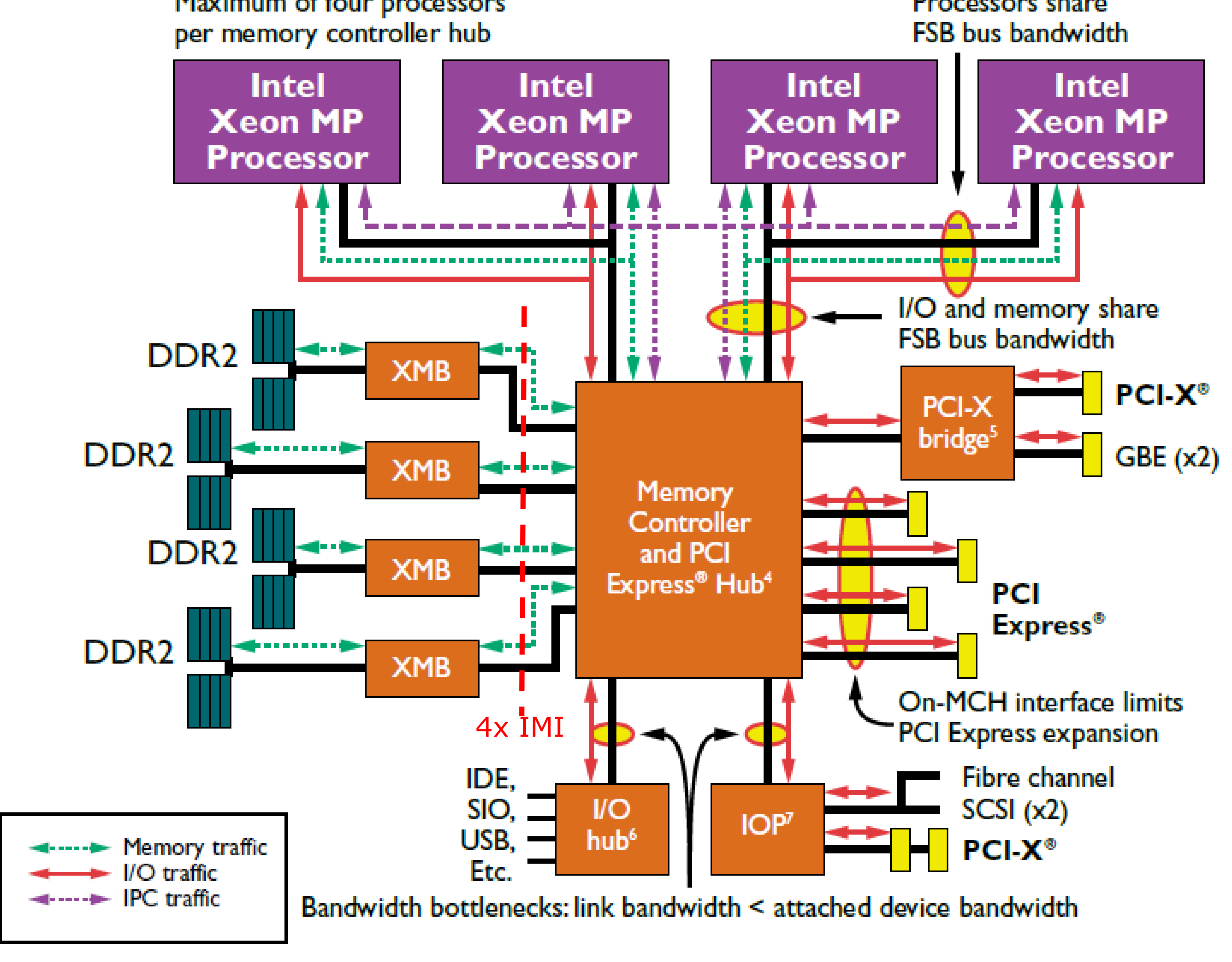
Intel's IMI (Independent Memory Interface)

Used to attach MBs to the MCH in the Pentium 4 Xeon based Truland MP platform (2005/2006) (for 1C/2x1C/2C)

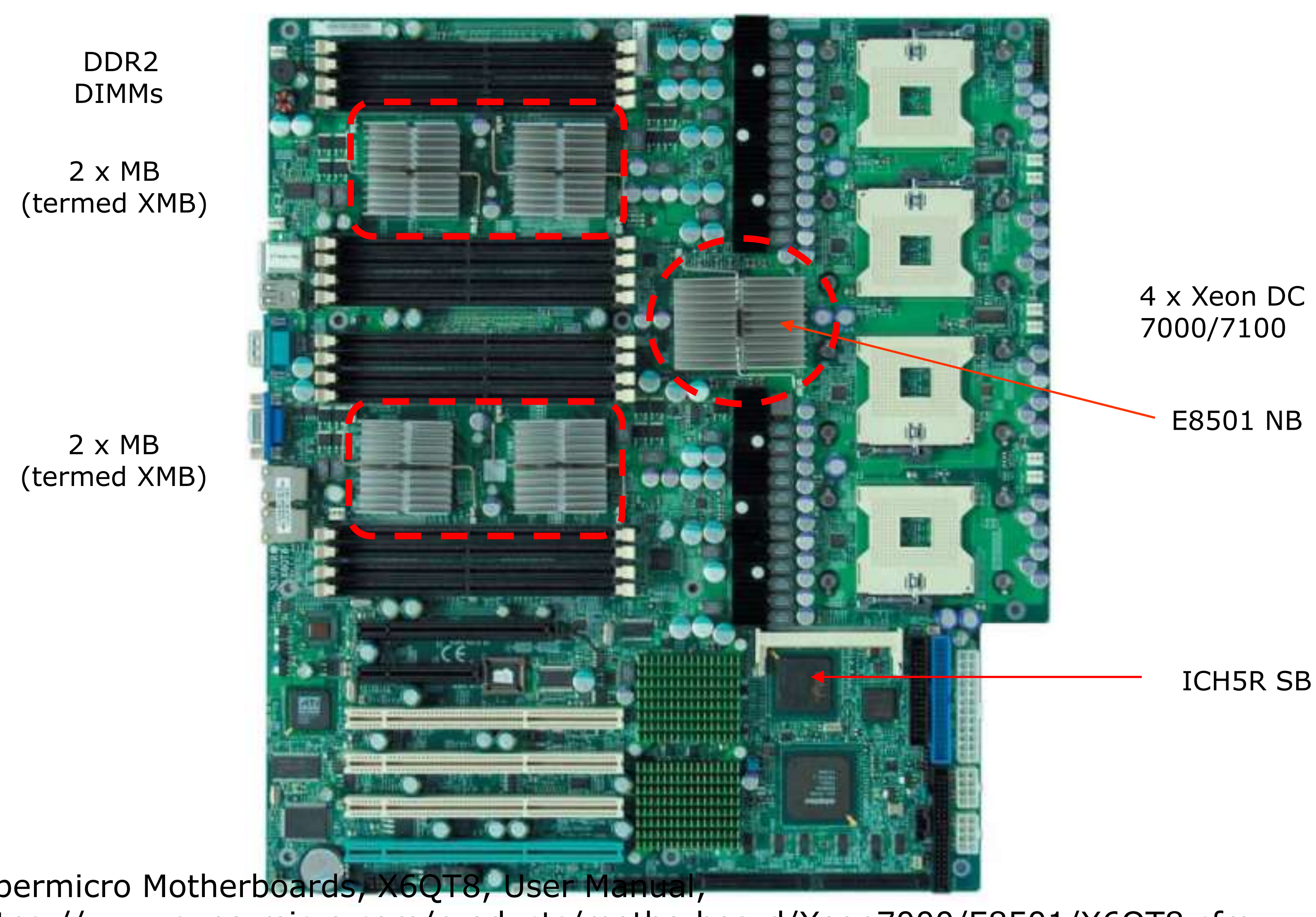
Using on-DIMM MBs (FB-DIMMs)

*

Example 1: Using low-line-count serial links and on-board MBs to attach DIMMs to the MCH in Intel's Pentium 4 based Truland MP server platform (2005) [164]



Example 1: Using low-line-count serial links and on-board MBs to attach DIMMs to the MC in Supermicro's Pentium 4 based Truland MP server board (X6QT8) (2006) [165



Supermicro Motherboards, X6QT8, User Manual,
https://www.supermicro.com/products/motherboard/Xeon7000/E8501/X6QT8.cfm

Using on-board MBs (on the mainboard/riser card -2

If using low-line-count serial links how to implement MBs?

Using on-board MBs (on the mainboard/riser card)

Using on-DIMM MBs (FB-DIMMs)

The MBs are placed on the mainboard or riser cards

Example 1:

Intel's IMI (Independent Memory Interface)

Used to attach MBs to the MCH
in the Pentium 4 Xeon based
Truland MP platform
(2005/2006)
(for 1C/2x1C/2C)

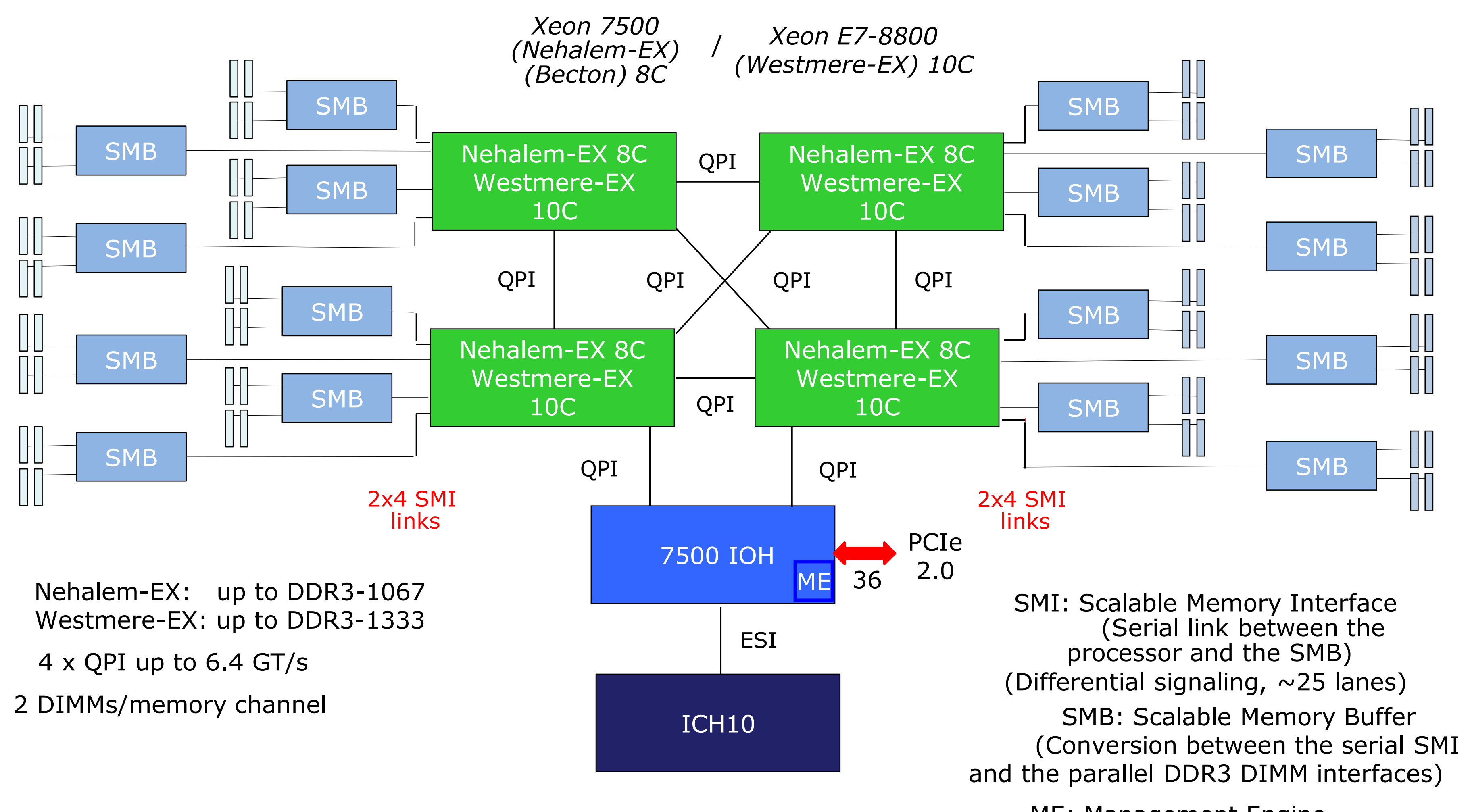
Example 2

Intel's SMI (Scalable Memory Interface)

Used to attach MBs to the processors
in the Nehalem-EX/Westmere-EX based
Boxboro-EX platform
(2010/2011)
(for 8C/10C)

(Introduced since no DDR3 based FB-DIMM memory became available)

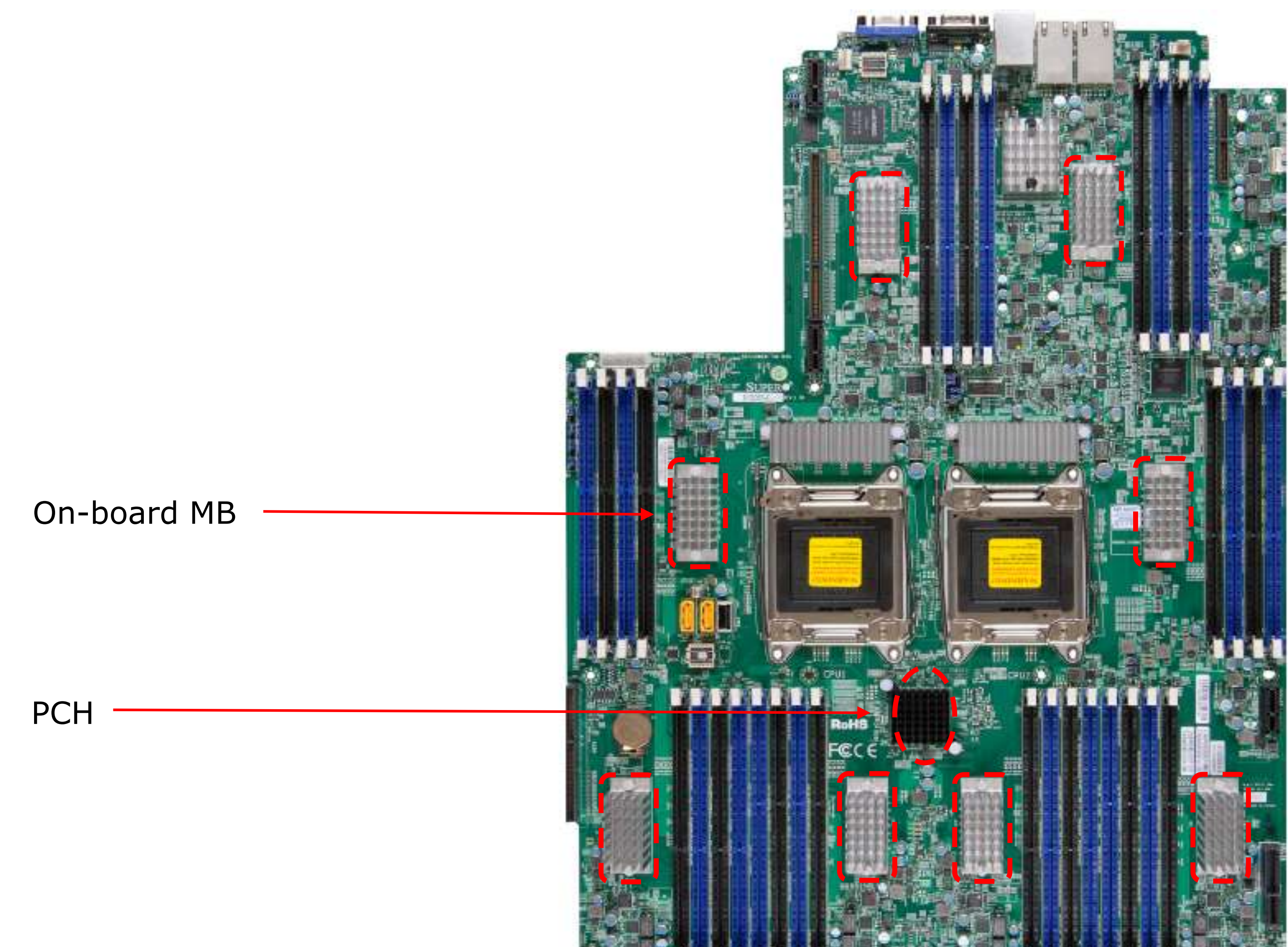
Example 2: Using low-line-count serial links and on-board MBs to connect DDR3 DIMMs to the processors in Intel's Boxboro-EX high-end 4S/8S server platform



Nehalem-EX/Westmere-EX based Boxboro-EX high-end 4S/8S server platform (for up to 10 C) (2010/2011)

2.4 The basic design space of connecting memory subsystems (17)

Example 2: Using low-line-count serial links and on-board MBs to connect DDR3 DIMMs to the processors in the Westmere-EX-based Boxboro-EX platform 2S Supermicro X10DBT server board) 2014) [167]



b2a2) Using on-DIMM MBs (FB-DIMMs)

If using low-line-count serial links how to implement MBs?

Using on-board MBs (on mainboard/riser card)

The MBs are placed on the mainboard or riser cards

Examples

Intel's IMI (Independent Memory Interface)

Used to attach MBs to the processors in the Pentium 4 Xeon based Truland MP platform (2005/2006) (for 1C/2x1C/2C)

Intel's SMI (Scalable Memory Interface)

Used to attach MBs to the processors in the Ivy Bridge/Nehalem-EX/Westmere-EX based Boxboro-EX platform (2010/2011) (for 8C/10C)

(*Introduced since* no DDR3 based FB-DIMMsy became available)

Using on-DIMM MBs (FB-DIMMs)

The MBs are placed directly on the DIMMs in form of FB-DIMMs

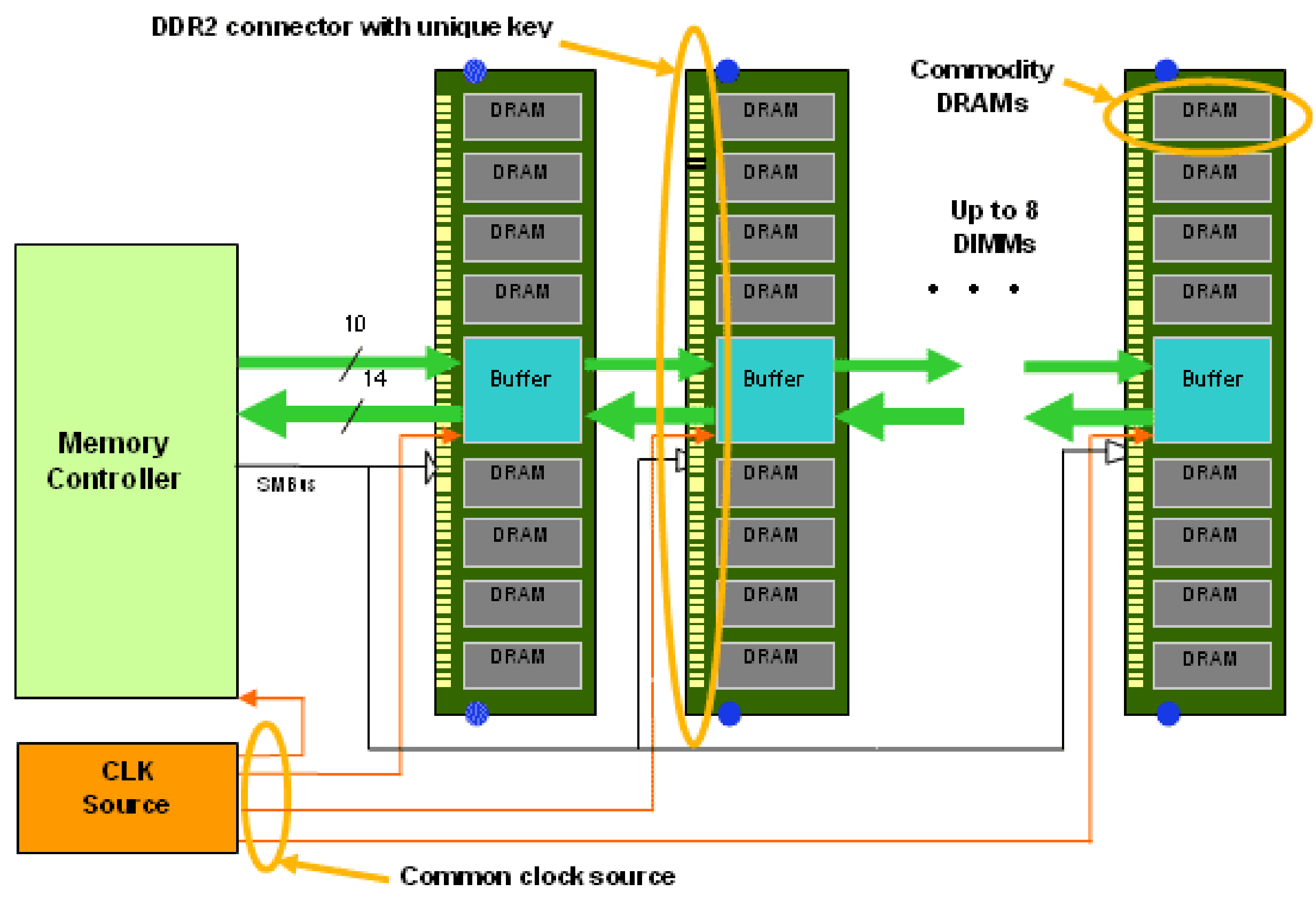
DDR2 based FB-DIMM interface

Used to attach DDR2 based FB-DIMM memory to the Core 2 and Penryn based Caneland platform (2007/2008)

(for 2C/2x2C/6C)

*

Example for connecting DIMMs via serial links and on-DIMM MBs (FB-DIMMs) to the MCs [166]



A brief introduction to FB-DIMM memories

- FB-DIMM memories were standardized and introduced in the 2006/2007 timeframe.
- They supported DDR2 memories up to DDR2-667.
- FB-DIMMs connect the MC and the DIMMs via low line count serial buses whereby the memory buffers providing serial/parallel conversion to the DIMMs are placed onto the DIMMs, as shown in the next Figure.

*

Layout of an FB-DIMM based memory subsystem -1 [102]

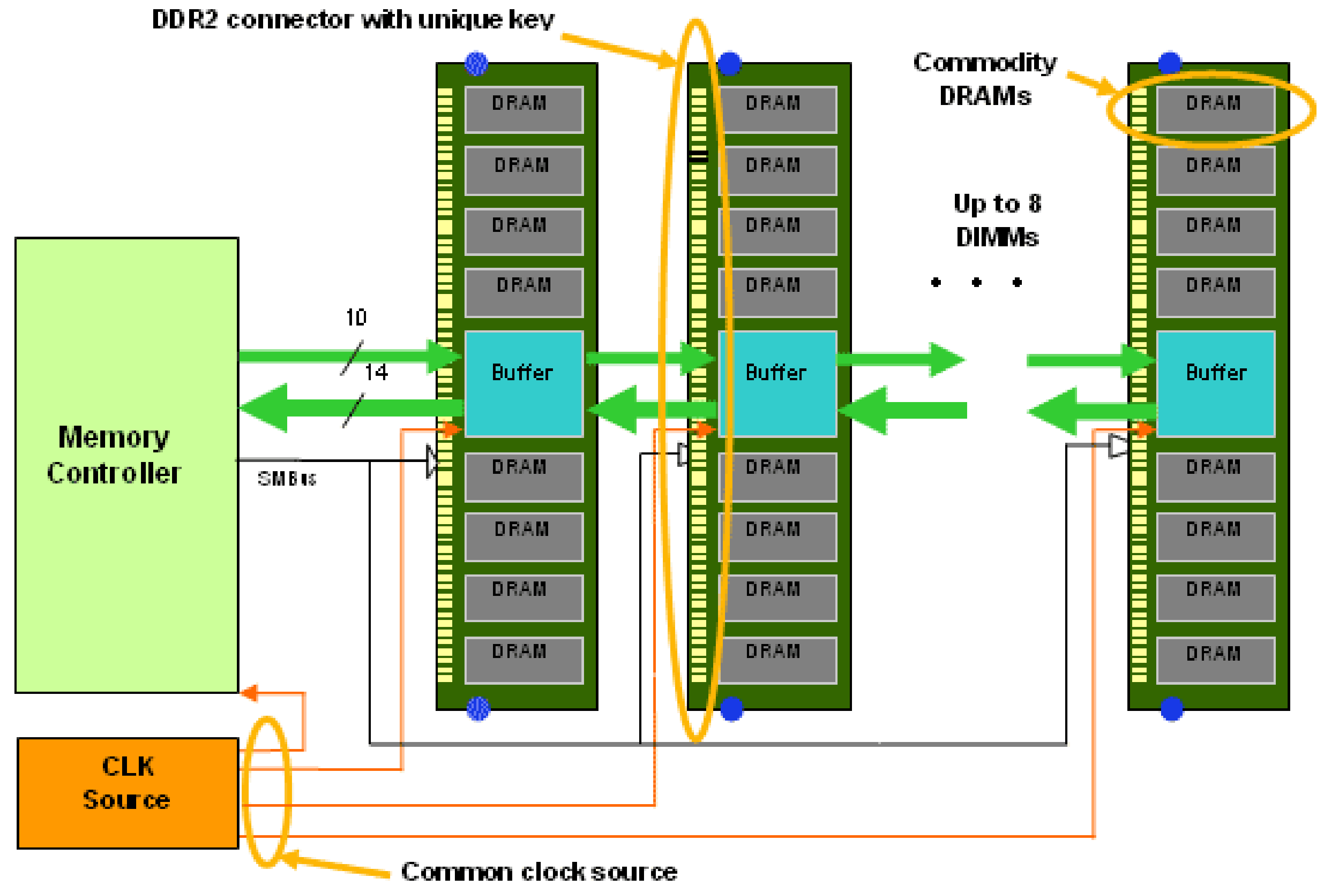


Figure: FB-DIMM memory architecture [102]

Layout of an FB-DIMM based memory subsystem -1 [102]

- The standardized FB-DIMM technology supports a cascaded connection of up to 8 FB-DIMMs..

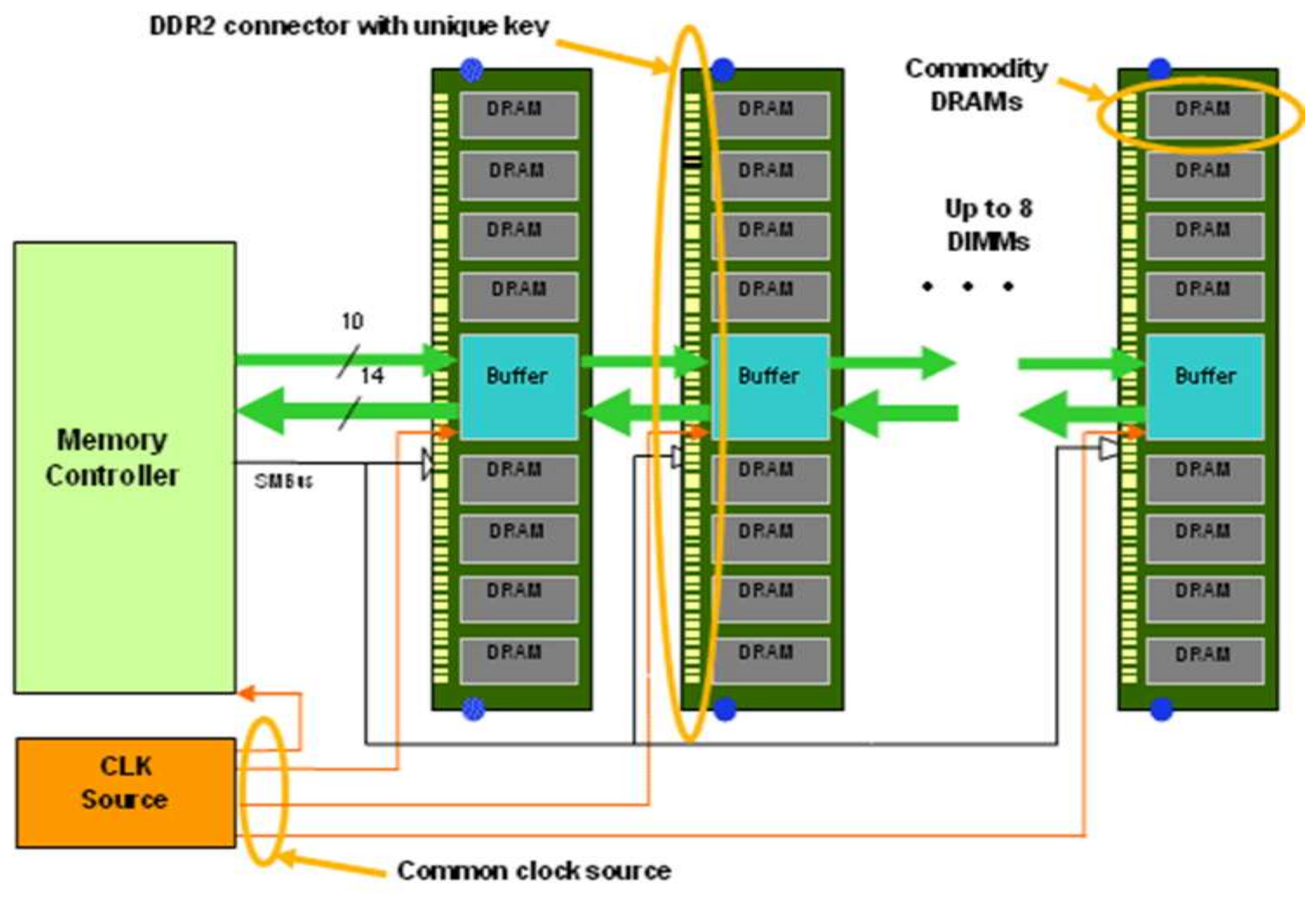


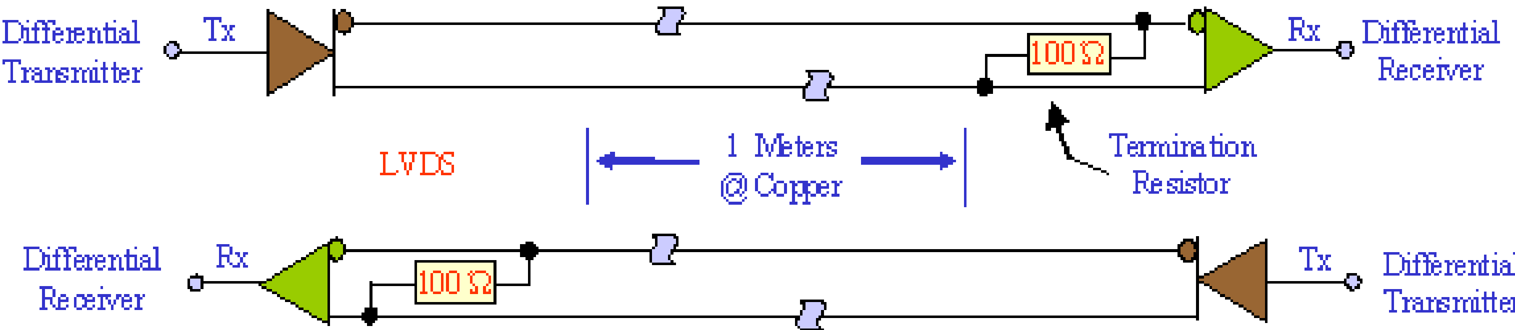
Figure: FB-DIMM memory architecture [102]

Layout of an FB-DIMM based memory subsystem -2 [102]

- The FB-DIMM technology implements packet based, bidirectional, point-to-point links.
- The links make use of differential signaling and includes 14 read/10 write lanes, as indicated in the next Figure.

*

Principle of differential signaling [170]



Layout of an FB-DIMM based memory subsystem -3 [102]

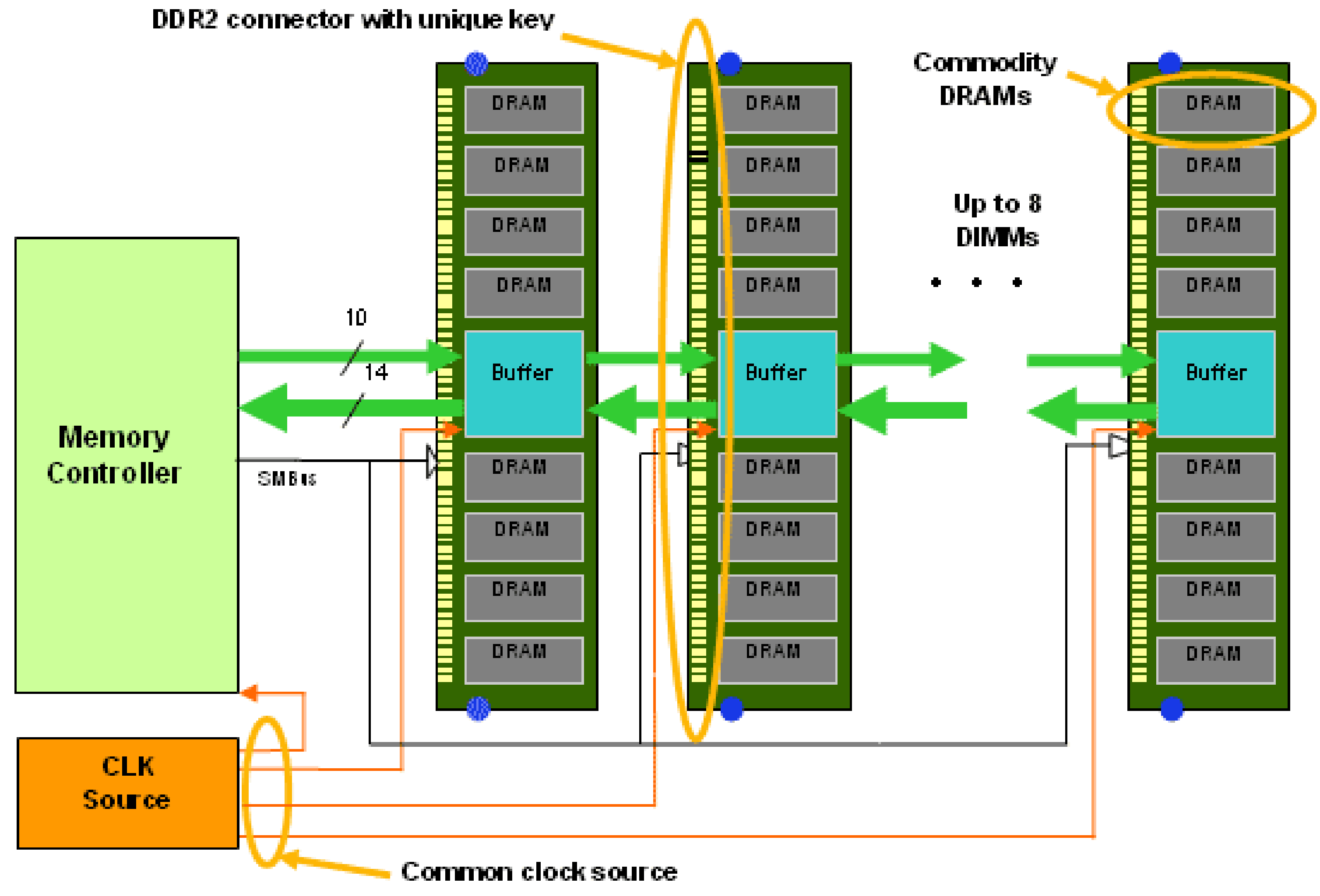


Figure: FB-DIMM memory architecture [102]

Layout of an FB-DIMM based memory subsystem -4 [103]

- The 14 inbound lanes carry data from the memory to the memory controller whereas the 10 outbound lanes transfer commands and data from the memory controller to the memory.
- The transfer rate is 6 times the transfer rate of the memory, i.e. 6x667=4 MT/s over the differential lines.
- Data and commands, transferred in 12 cycles, from one frame, that is an inbound frame includes 12x14=168 bit whereas an outbound frame has 12x10=120 bits.
- For more information we refer to the literature, e.g. [103].

Remarks

- FB-DIMMs have a number of inherent drawbacks, such as
  - higher dissipation due to the necessary serial/parallel conversions,
  - longer access times because of cascading DIMMs and
  - the resulting higher price.
- As a consequence, there was a small demand for DDR2-based FBDIMM memories and memory vendors were reluctant to develop DDR3-based FB-DIMMs.
- Thus, after the emergence of DDR3 memories Intel was forced to move back to custom DDR3 DIMMs connected by serial links and on-board memory buffers to the MCs in their subsequent Boxboro-EX platform.

Introducing and cancellation of using on-DIMM MBs (FB-DIMMs)

If using low-line-count serial links how to implement MBs?

Using on-board MBs (on mainboard/riser card)

The MBs are placed on the mainboard or riser cards

Examples

Intel's IMI (Independent Memory Interface)

Used to attach MBs to the processors in the Pentium 4 Xeon based Truland MP platform (2005/2006) (for 1C/2x1C/2C)

Intel's SMI (Scalable Memory Interface)

Used to attach MBs to the processors in the Ivy Bridge/Nehalem-EX/Westmere-EX based Boxboro-EX platform (2010/2011) (for 8C/10C)

(Introduced since no DDR3 based FB-DIMMsy became available)

Using on-DIMM MBs (FB-DIMMs)

The MBs are placed directly on the DIMMs in form of FB-DIMMs

DDR2 based FB-DIMM interface

Used to attach DDR2 based FB-DIMM memory to the MCH in the Core 2 and Penryn based Caneland platform

after DDR2 based FB-DIMMs appeared.

(2007/2008)

(for 2C/2x2C/6C)

*

b2b) Using low-line-count parallel links to connect MBs to the MCs

Connecting DIMMs via low-line-count links and memory buffers (MBs) to the MCs

Using low-line-count serial links to connect MBs to the MCs

The links are implemented as two unidirectional, Point-to-Point serial, packet based channels, typically with differential signaling.

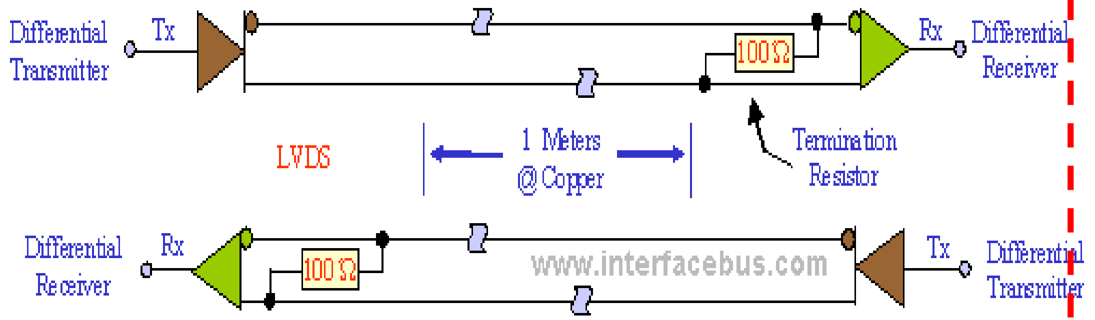


Figure: Example for differential signaling [170]

Drawback: The MBs have to perform serial/parallel conversion, this increases dissipation.

Used in Intel's early 4S or 8S Truland MP and Boxboro-EX platforms (2005 - 2011) for 2C to 10C

Implemented links (IMI, SMI) have about 70 lines (vs. 240 for DDR2/3).

Using low-line-count parallel links to connect MBs to the MCs

The links are implemented as 64-bit parallel, bidirectional channels typically with single-ended voltage reference signaling, termed VMSE (Voltage Mode Single Ended) signaling

Benefit: No need for serial/parallel conversion in the MBs, only interface conversion is required

Used in Intel's Brickland platform (2014 -2016) for 15C/18C/24C

Intel's implementation (the SMI 2 interface) comprises about 110 lines.

If using low-line-count parallel links how to implement MBs?

Using low-line-count parallel links to connect MBs to the MCs

Using on-board MBs
(on mainboard/riser card)

There is a need for an interface converter, called Memory Buffer (MB), between the low-line-count parallel interface and the standard DDR interface.

The MBs are placed either onto the mainboard or riser cards.

Example

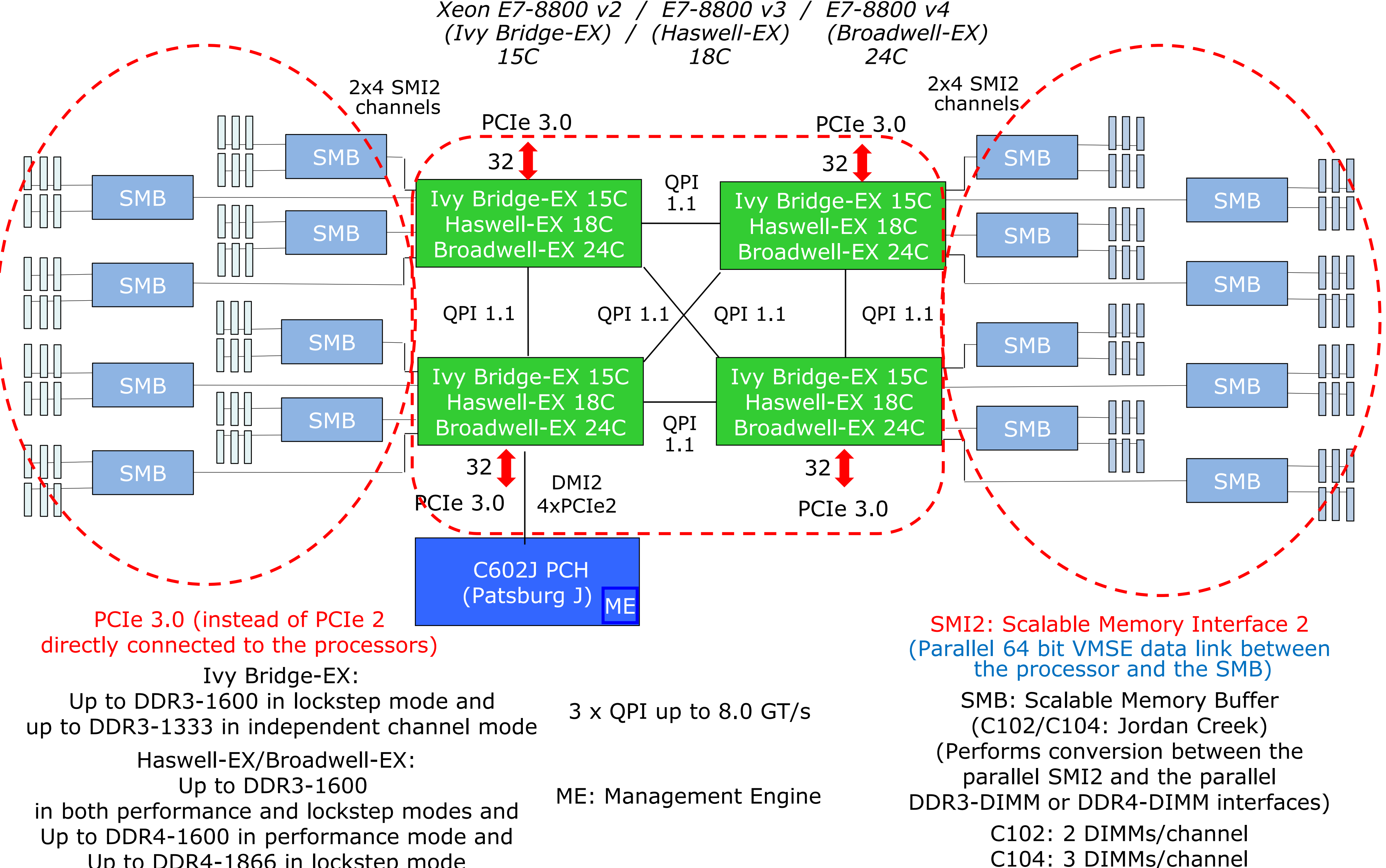
Intel's SMI 2 (Scalable Memory Interface 2 in the Brickland platform that targets the Ivy Bridge-EX/Haswell-EX/Broadwell-EX processors.

(2014-2016)

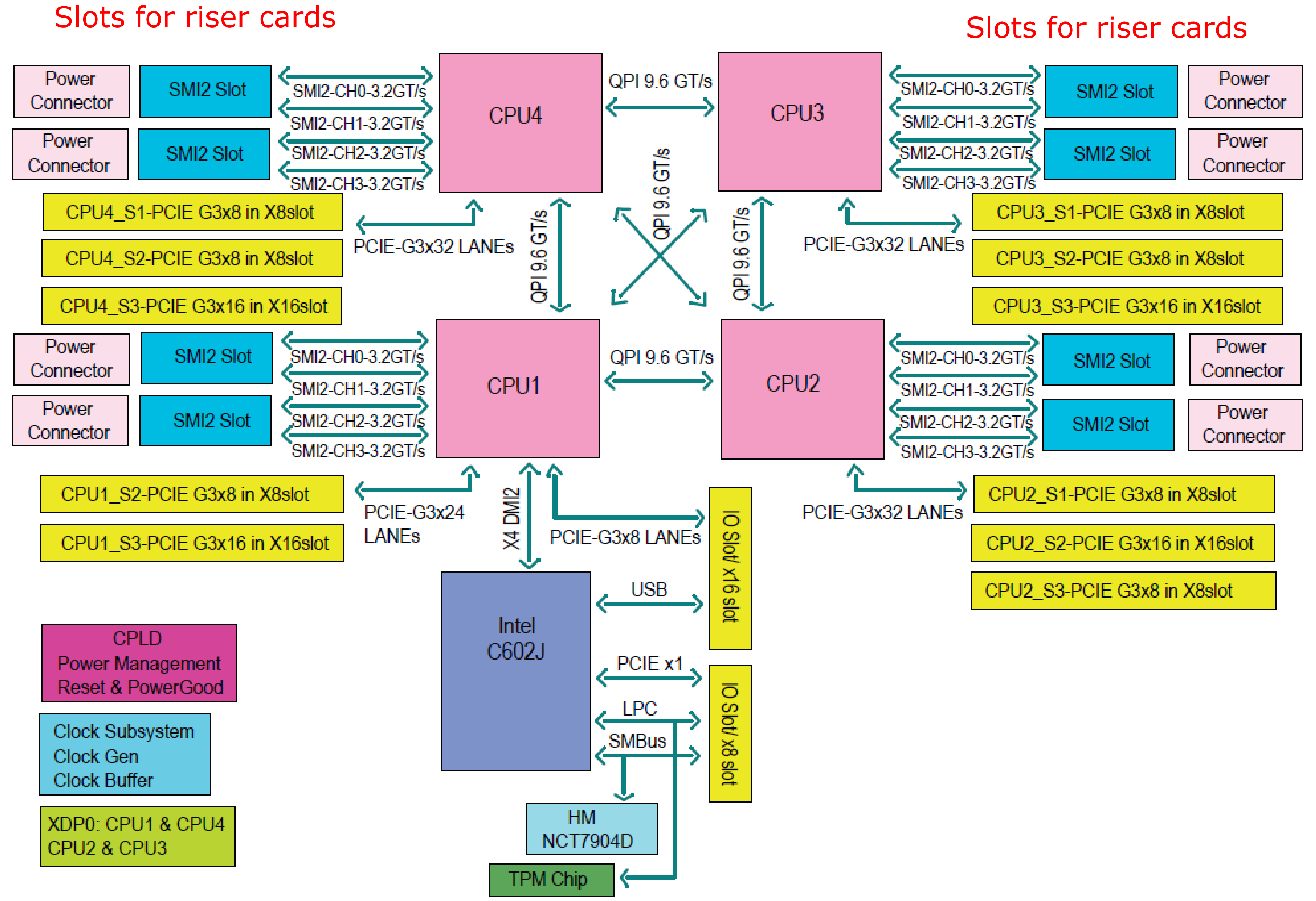
(for up to 15C/18C/24C)

*

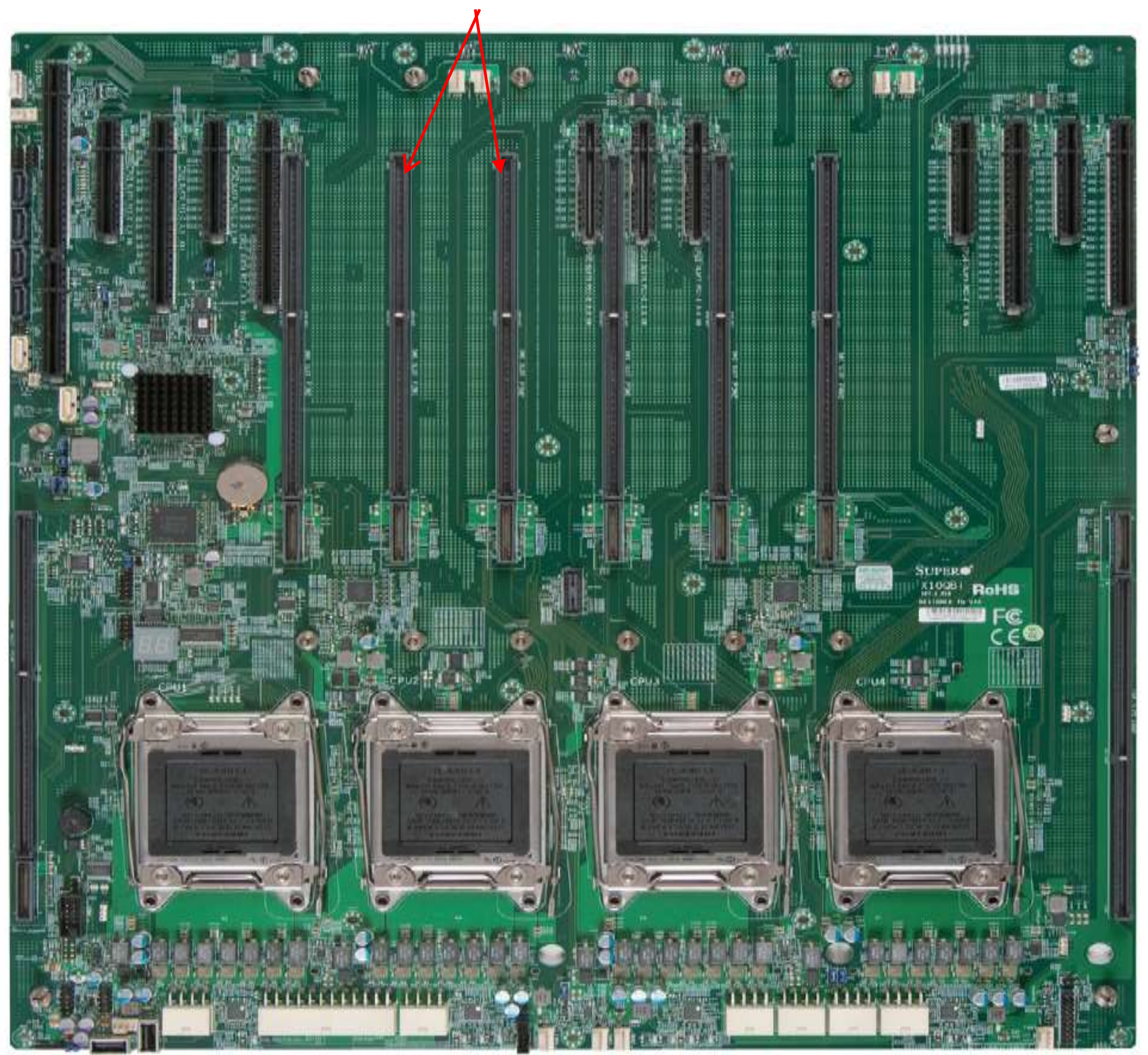
Example: Using low-line-count parallel links and on-board MBs to connect DDR3/4 DIMMs to the processors in Intel's Brickland high-end 4S/8S server platform -1



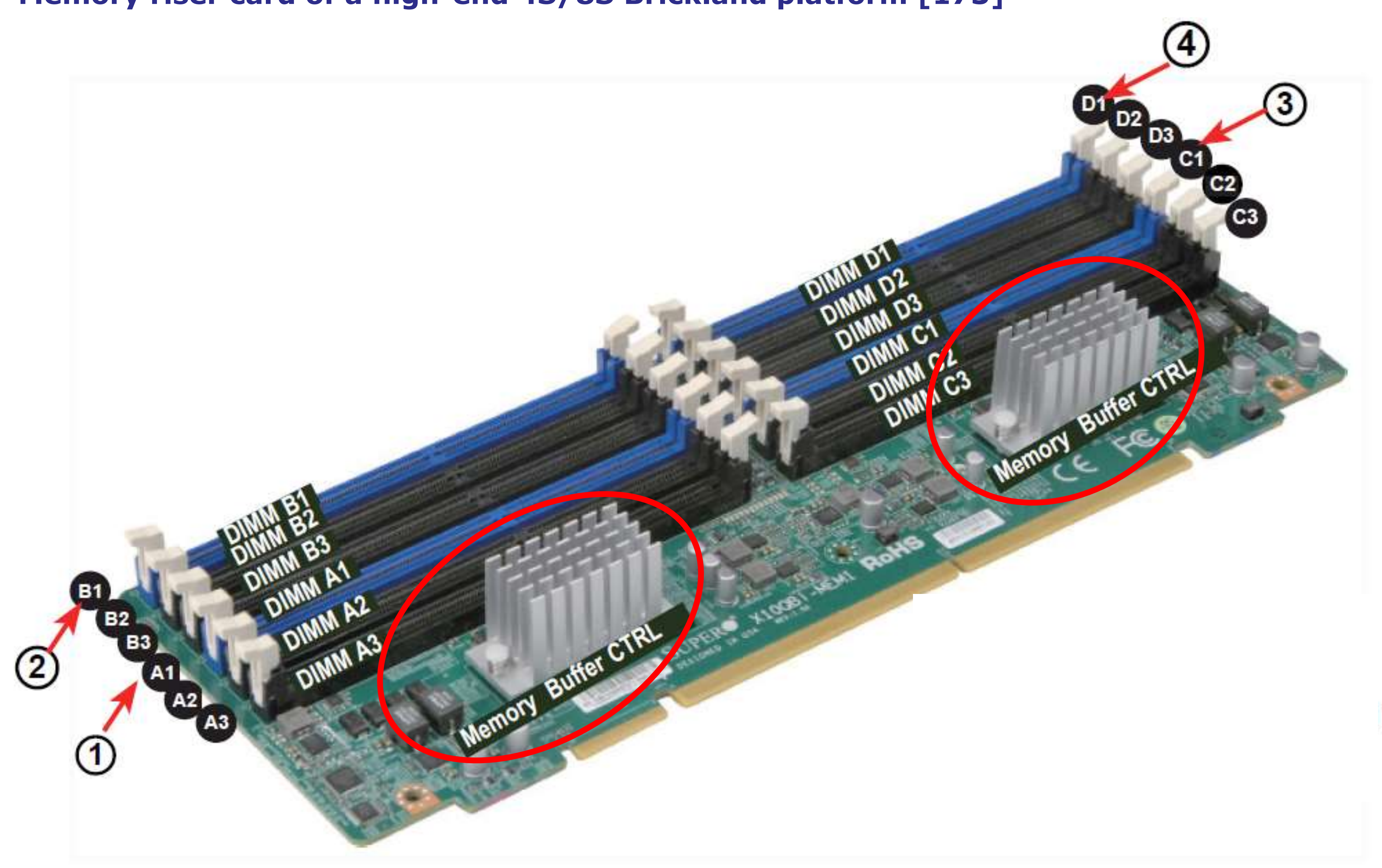
Block diagram of a high-end 4S/8S Brickland platform based server using riser cards [173]



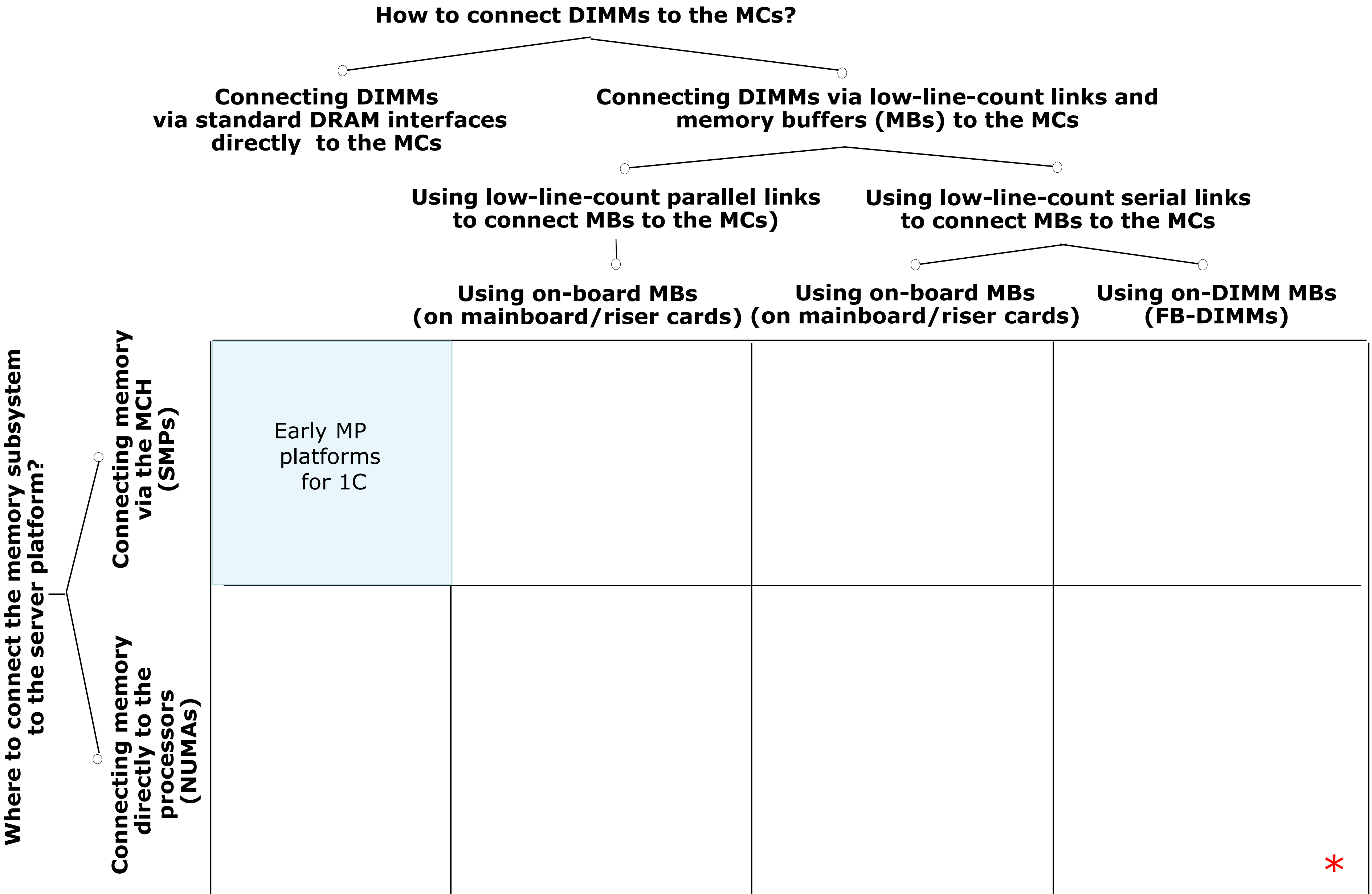
Mainboard of a high-end 4S/8S Brickland platform based server using riser cards [173]



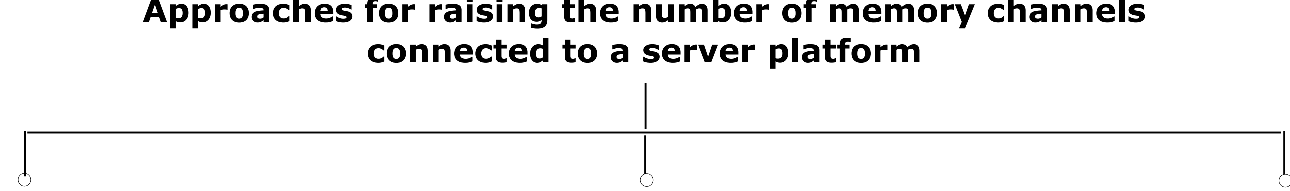
Memory riser card of a high-end 4S/8S Brickland platform [173]



c) Design space of connecting memory subsystems to platforms -1

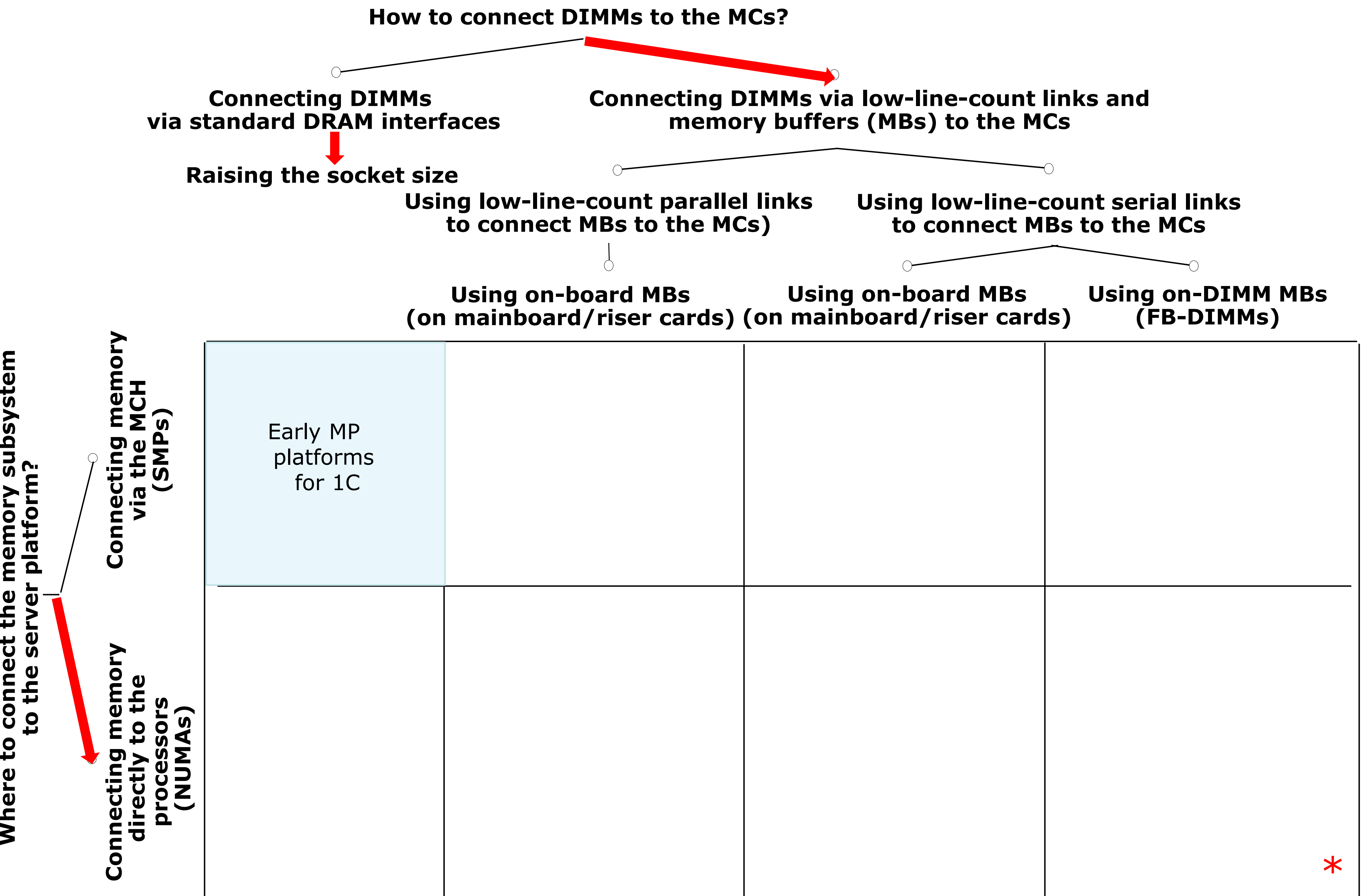


Approaches for raising the number of memory channels connected to a server platform -1

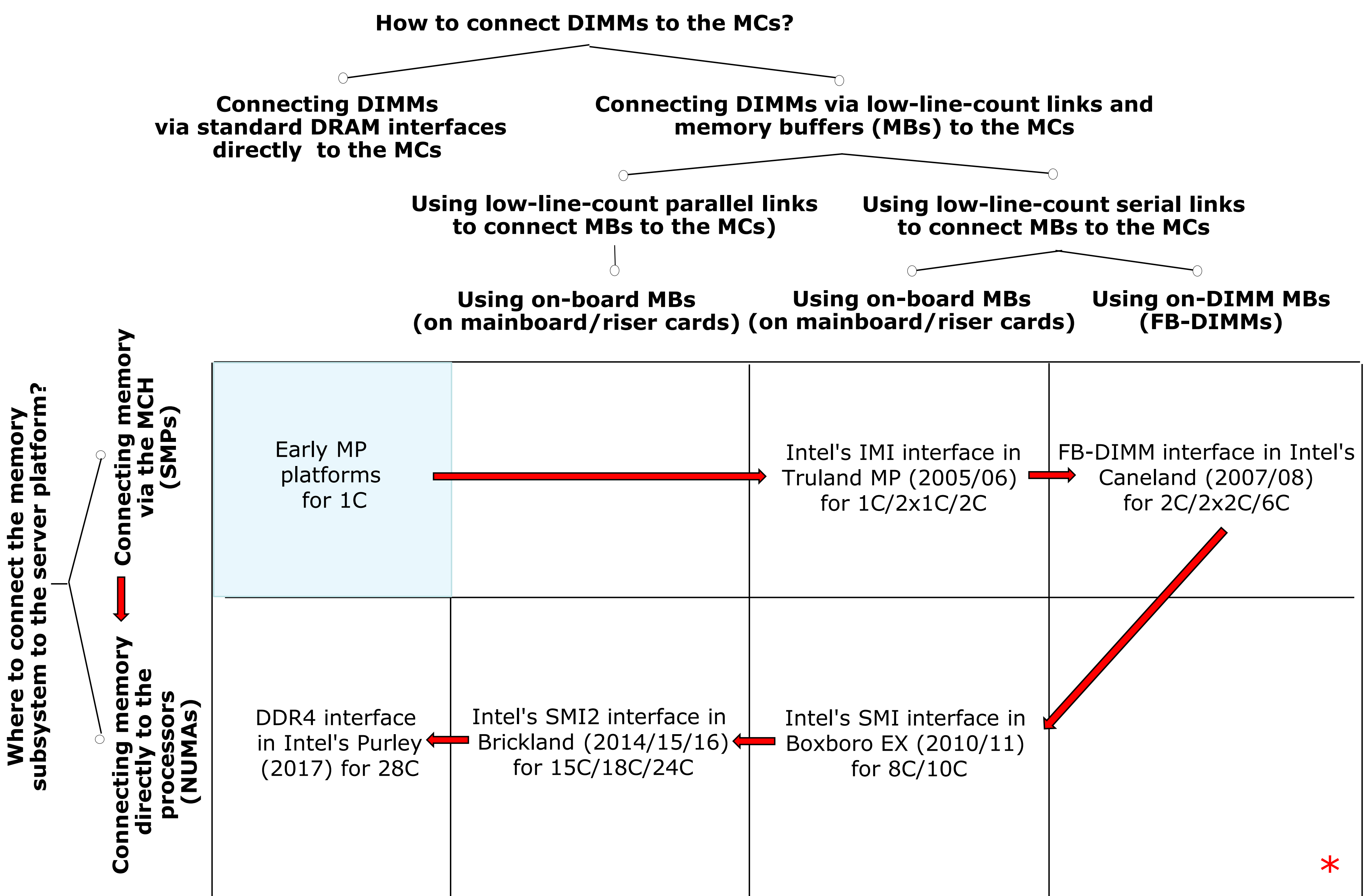


The next Figure shows how these approaches relate to the design space of connecting memory subsystems to platforms.

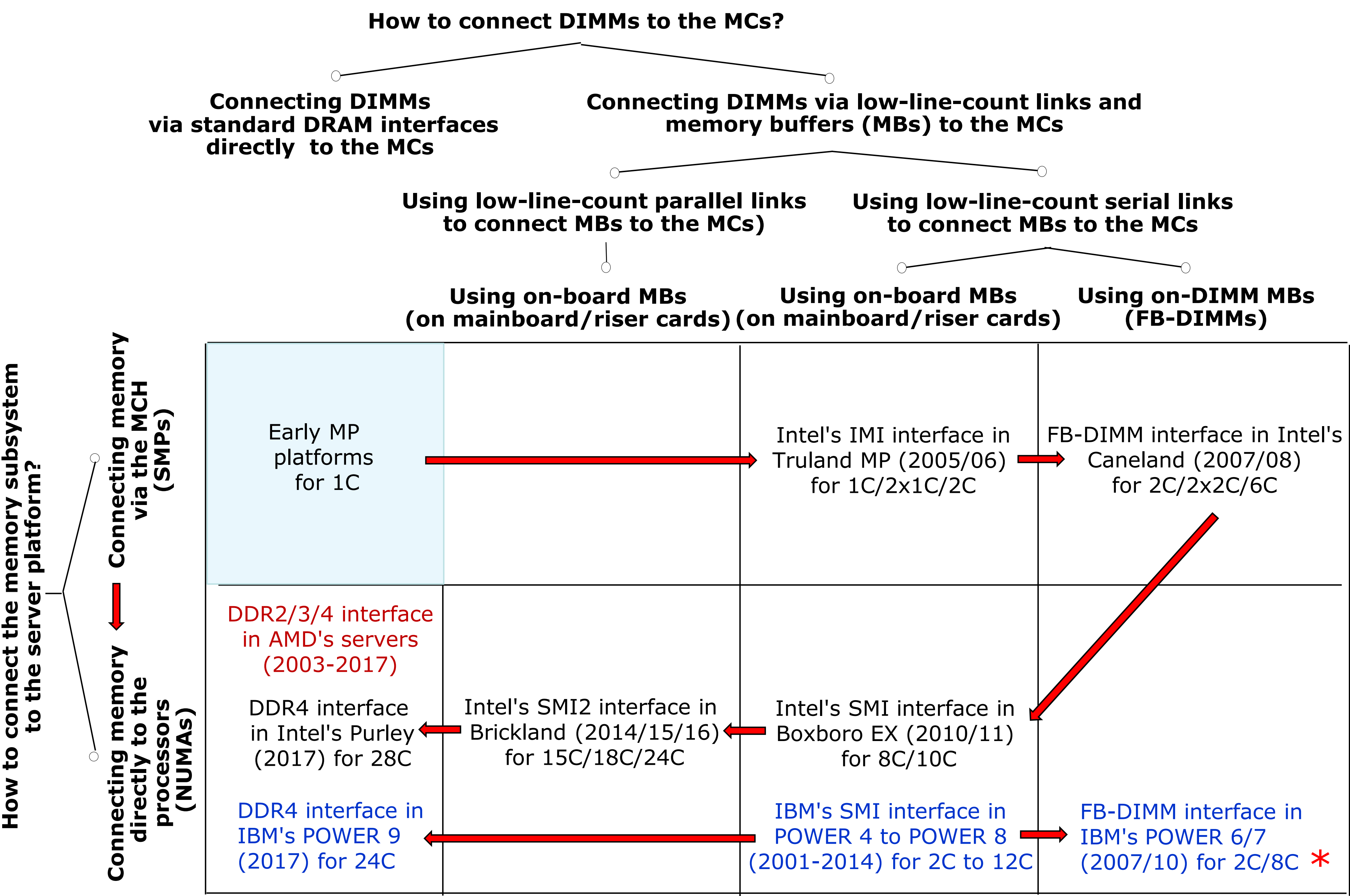
Approaches for raising the number of memory channels connected to a server platform -2

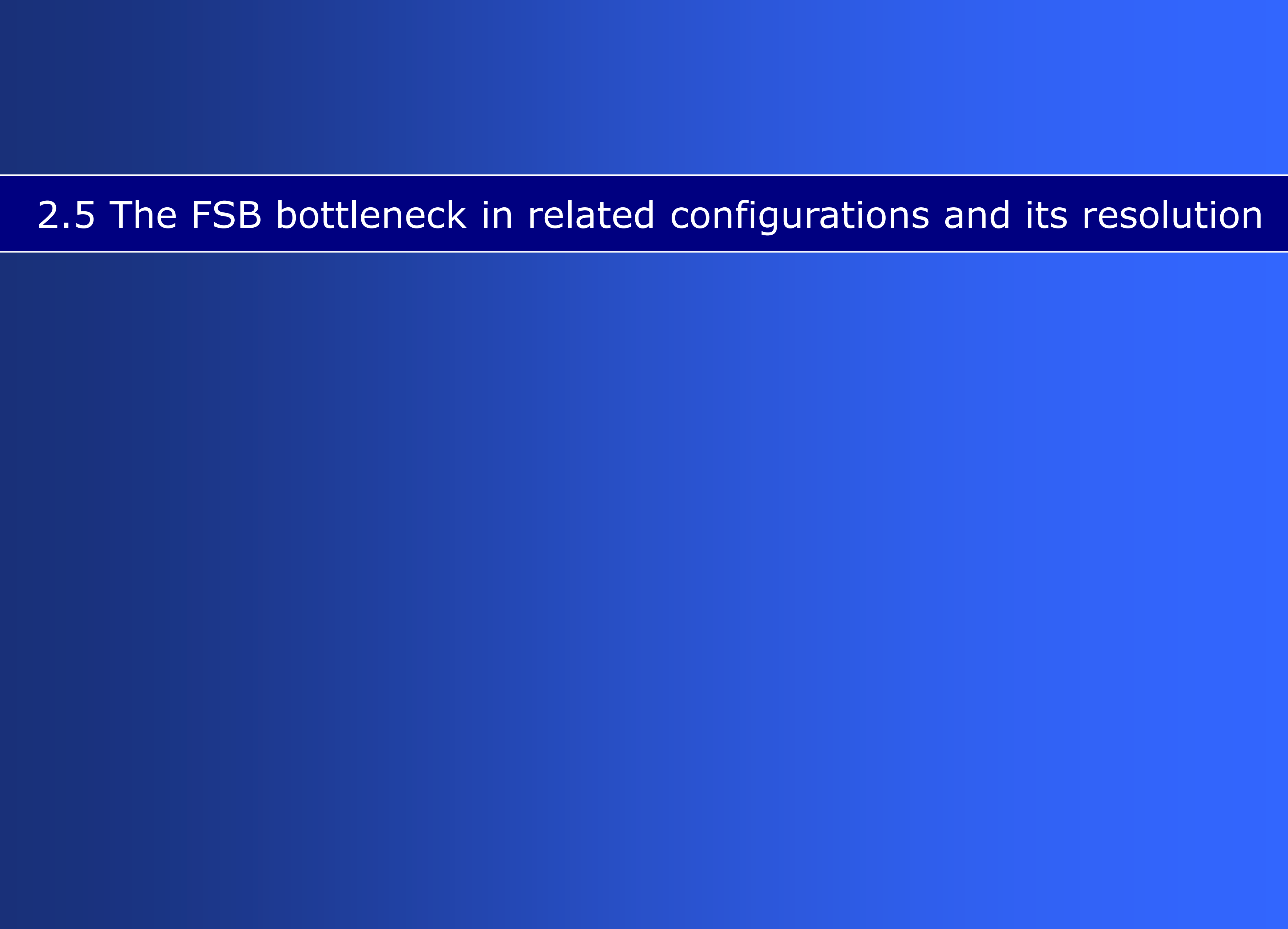


Evolution of connecting the memory subsystem in Intel's high-end server platforms



Evolution of connecting the memory subsystem in AMD's and IBM's server platforms




2.5 The FSB bottleneck in related configurations and its resolution

2.5 The FSB bottleneck in SMP configurations and its resolution

In early SMP configurations (e.g. in first Pentium 4 based single core (SC) Xeon based systems) there is a single FSB that interconnects the four processors with the NB (Northbridge), as seen.

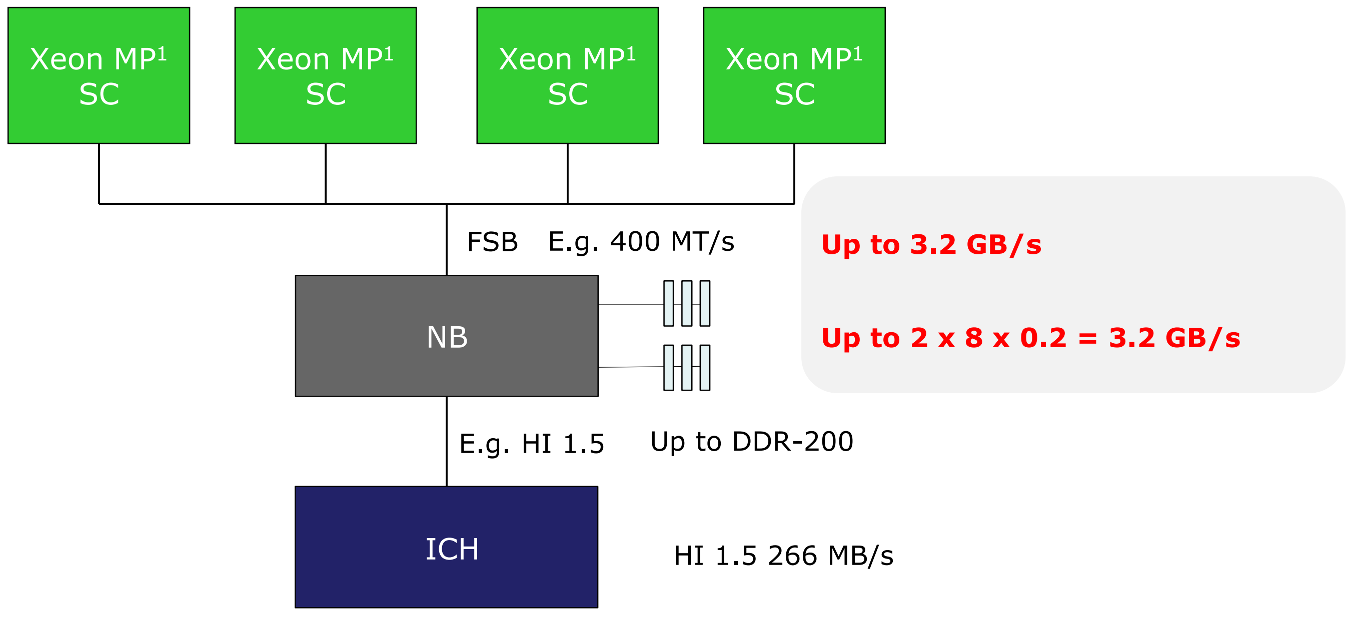


Figure: Block diagram of a single core Pentium 4 MP server platform (Based on [101])

Note that the FSB and the memory subsystem have the same maximal bandwidth.

Remark

In the referenced platform the memory subsystem is interconnected in a different way as indicated in the Figure by implementing 4x16-bit data and a separate address link.

*

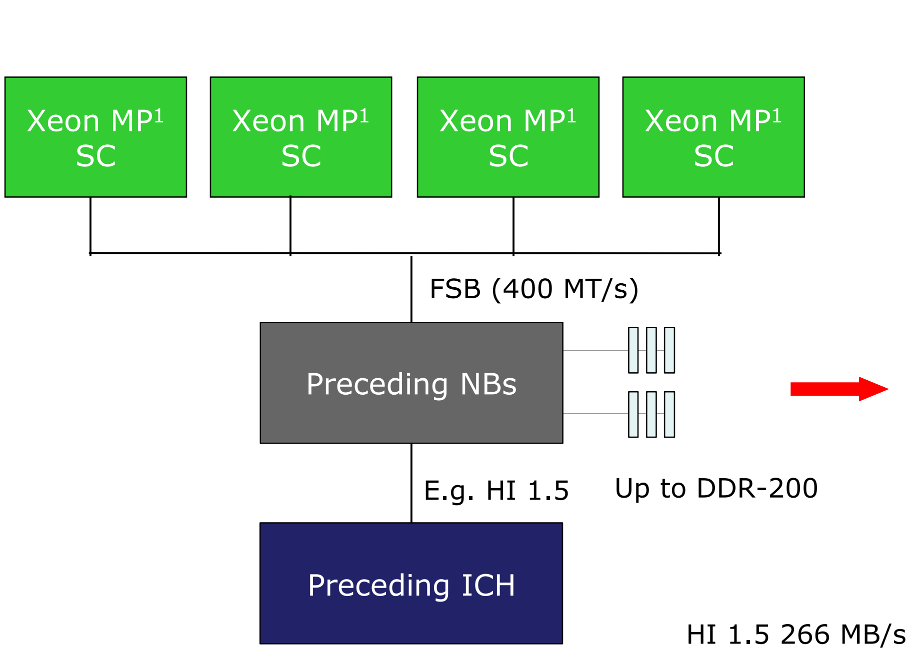
Resolving the FSB bottleneck in SMP configurations

Obviously, in SMP configurations with higher core count and memory speeds, a single FSB may severely limit memory bandwidth and thus performance.

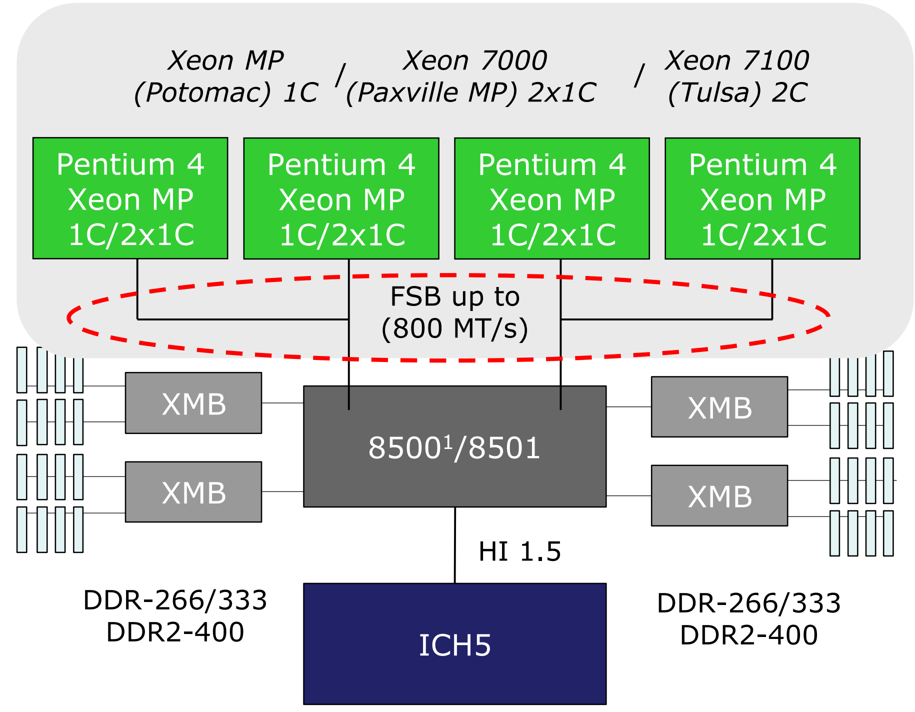
This FSB caused bandwidth bottleneck may resolved by using more than one FSB, as indicated in the subsequent Figures.

*

Resolving the FSB bottleneck in dual core 4S server platforms by using dual FSBs



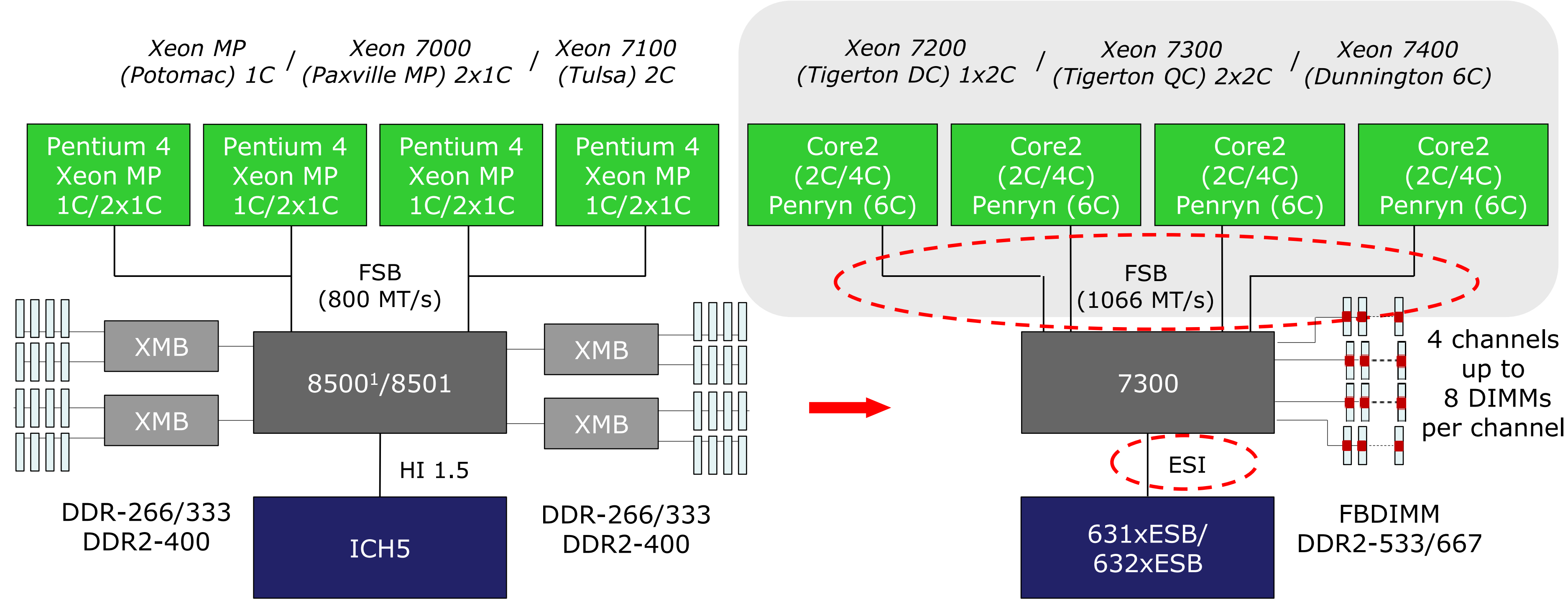
Previous Pentium 4 MP aimed
MP server platform (for single core processors)
(2004 and before)



90 nm Pentium 4 Prescott MP aimed
Truland MP server platform (for up to 2 C)
(2005)

[1]The 8500 MCH support only667 MT/sFSB speeds and is used in single core configurations whereas the 8501 supports already 800 MT/s and is used for dual core configurations

Resolving the FSB bottleneck in up to 6 core 4S server platforms by using quad FSBs



90 nm Pentium 4 Prescott MP aimed
Truland MP server platform (for up to 2 C)
(2006)

HI 1.5 (Hub Interface 1.5)
8 bit wide, 66 MHz clock, QDR,
266 MB/s peak transfer rate

Core 2 aimed
Caneland MP server platform (for up to 6 C)
(2007)

ESI: Enterprise System Interface
4 PCIe lanes, 0.25 GB/s per lane (like the DMI interface,
providing 1 GB/s transfer rate in each direction)

[1] The E8500 MCH supports an FSB of 667 MT/s and consequently only the SC Xeon MP (Potomac)

2.6 Overview of the evolution of Intel's high-end 4S/8S server platforms

2.6 Overview of the evolution of Intel's high-end 4S/8S server platforms

Platform	Platform topology	Date of intro.	Processor	Techno-logy	Core count (up to)	No./kind of the links to mem. buffers	No. of mem. channels/ mem. buffer	No. /speed of mem. channels (up to)
Truland MP	SMP w/dual FSBs	3/2005	90 nm Pentium 4 MP (Potomac)	90 nm	1C	4 serial low-line-count links (IMI) between the MCH and mem. buffers (XMBs)	2 memory channels/ XMB	2x DDR2-800/ socket
		11/2005	Xeon 7000 (Paxville MP)	90 nm	2x1C			
		8/2006	Xeon 7100 (Tulsa)	65 nm	2C			
Caneland	SMP w/Quad FSBs	9/2007	Xeon 7200 Tigerton DC (Core 2)	65 nm	2C	4 serial low-line-count FB-DIMM links between the MCH and mem. buffers (AMBs)	1 memory channel/ AMB	1x DDR2-667/ socket
		9/2007	Xeon 7300 Tigerton QC (Core 2)	65 nm	2x2C			
		9/2008	Xeon 7400 (Dunnington) (Penryn)	45 nm	6C			
Boxboro-EX	NUMA fully connect. by QPI buses	3/2010	Nehalem-EX (Xeon 7500/ (Beckton)	45 nm	8C	4 serial low-line-count links (SMI) per socket to 4 mem. channels (SMBs)	2 memory channels/ SMB	8x DDR3-1067/ socket
		4/2011	Westmere-EX (E7-8800)	32 nm	10C			
Brickland	NUMA fully connect. by QPI buses	2/2014	Ivy Bridge-EX (E7-8800 v2)	22 nm	15C	4 parallel low-line-count links (SMI2) per socket to 4 mem. buffers (SMBs)	2 memory channels/ SMB	8x DDR3-1600 (IWB) / DDR4-1866 (HSW)/ socket
		5/2015	Haswell-EX (E7-8800 v3)	22 nm	18C			
		6/2016	Broadwell-EX (E7-8800 v4)	14 nm	24C			
Purley	NUMA fully connect. by UPI buses	7/2017	Skylake-SP	14 nm	28C	6 psrallel DDR4 channels per socket	Direct attached	6x DDR4-2666 mem. speed

2.6 Overview of the evolution of Intel's high-end 4S/8S server platforms (2)

a) From single core to dual core high-end 4S server platforms (called the Truland MP server platform) -1

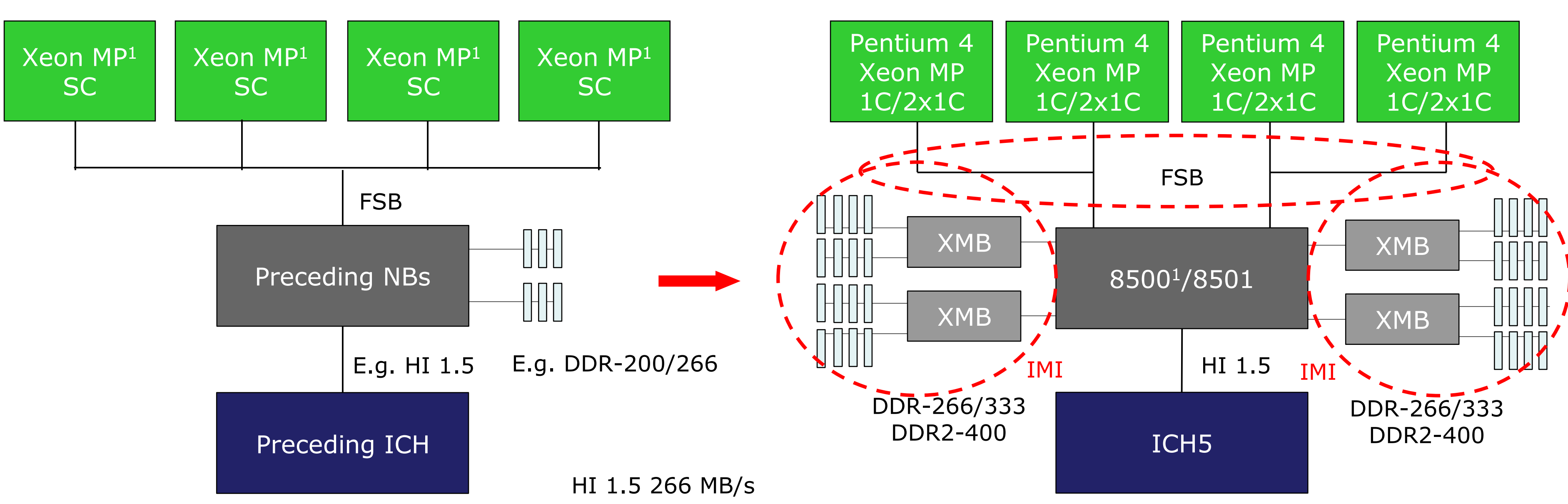
Aims

a) Providing more memory channels per processor by introducing low-line-count serial links and memory buffers (placed onto the motherboard) for interconnecting the memory controller that is implemented in the MCH and the DIMMs.

b) Providing dual FSBs to avoid FSB bottleneck.

*

a) Evolution of single core high-end 4S server platforms to dual core high-end 4S server platforms (called Truland MP server platform) -2



[1]Previous Pentium 4 MP aimed
MP server platform (for single core processors)
(2004 and before)

90 nm Pentium 4 Prescott MP aimed
Truland MP server platform (for up to 2 C)
(2006)

IMI (Independent Memory Interface): Low-line-count (70 signal lines) serial interface vs. DDR2 with 240 lines.
XMB: eXternal Memory Bridge

*

b) Evolution from dual core to up to 6-core high-end 4S server platforms (called the Caneland server platform) -1

Aims

a) Utilizing (DDR2 based) FB-DIMM memory with serial, low-line-count links to the memory controller (which is implemented in the MCH) and on-DIMM memory buffers in-order to simplify system design.

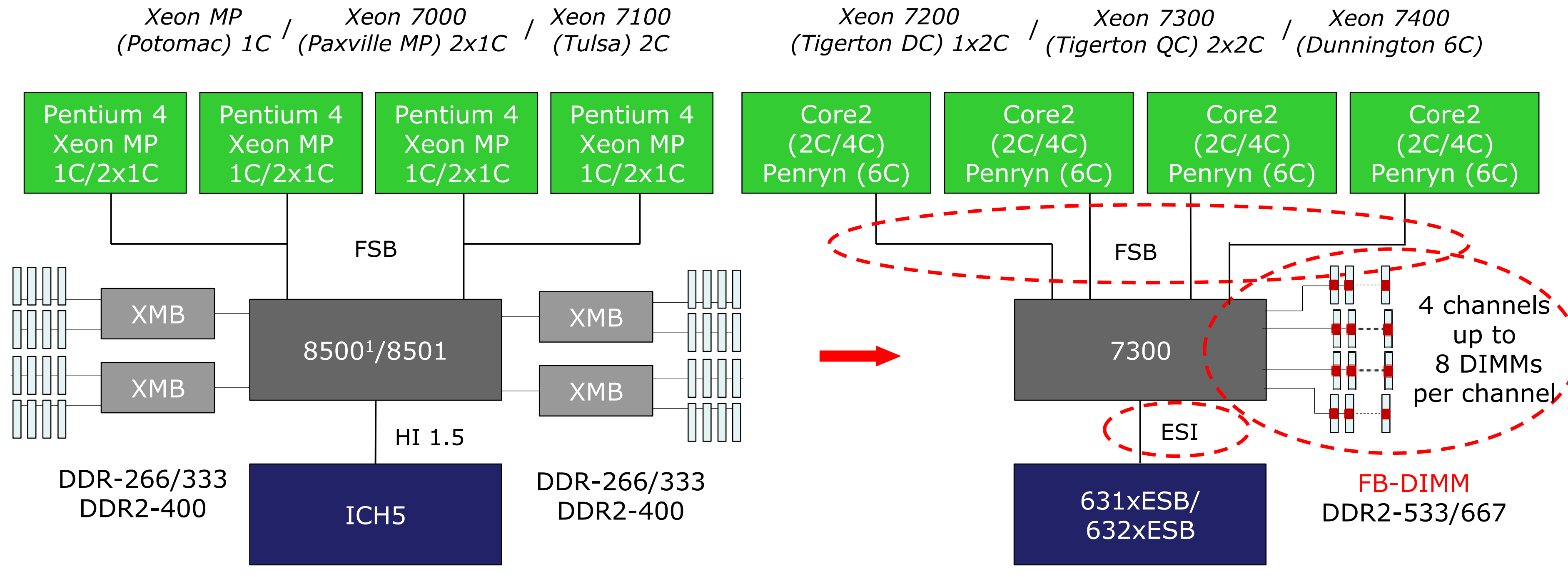
b) A further objective is to raise FSB bandwidth by providing 4 FSB links to the processors.

Remarks

- FB-DIMMs need only a fraction of the lines compared to standard DDR2 DIMMs (about 50 vs. 240) thus significantly more memory channels (up to 6 channels) may be connected to the MCH than in case of directly connected high-line-count DDR2 DIMMs.
- Furthermore, due to the cascaded nature of interconnecting FB-DIMMs (with repeater functionality), up to 8 DIMMs may be placed into each DIMM channel instead of two or three as typical for standard DDR2 memory channels.
- This results in considerably higher memory bandwidth and memory size by reduced mainboard complexity.
- Based on these benefits Intel decided to use FB-DIMM-667 memory in their server platforms, first in their DP platforms already in 2006 followed by the Caneland MP platform in 2007, as indicated before.

*

b) Evolution of dual core high-end 4S server platforms to up to 6 core high-end 4S platforms (called Caneland server platform) -2



90 nm Pentium 4 Prescott MP aimed
Truland MP server platform (for up to 2 C)
(2006)

Core 2 aimed
Caneland MP server platform (for up to 6 C)
(2007)

HI 1.5 (Hub Interface 1.5)
8 bit wide, 66 MHz clock, QDR,
266 MB/s peak transfer rate

ESI: Enterprise System Interface
4 PCIe lanes, 0.25 GB/s per lane (like the DMI interface, providing 1 GB/s transfer rate in each direction)

[1] The E8500 MCH supports an FSB of 667 MT/s and consequently only the SC Xeon MP (Potomac)

*

c) Evolution to the up to 10-core Boxboro-EX 4S/8S high-end server platform (targeting the Nehalem-EX/Westmere-EX processors) -1

Aims

a) Raising per socket memory bandwidth of the 4S/8S server by connecting memory directly to the processors.

b) Making use of DDR3 memory.

- Since no DDR3-based FB-DIMMs became available Intel could use DDR3 memory only when it falls back on using custom DDR3 DIMMs and on-board memory buffers nevertheless this time connected to the processors rather than to the MCH.
- The reason for the cancellation of DDR3-based FB-DIMM developments was the small market share achieved by DDR2-based FB-DIMMs.

  This was caused by the inherent drawbacks of FB-DIMM memories, such as

  - higher dissipation due to the necessary serial/parallel conversions,
  - longer access times because of cascading DIMMs and
  - the resulting higher price.

Returning to on-board memory buffers using FB-DIMMs in the Boxboro-EX high-end platform -2

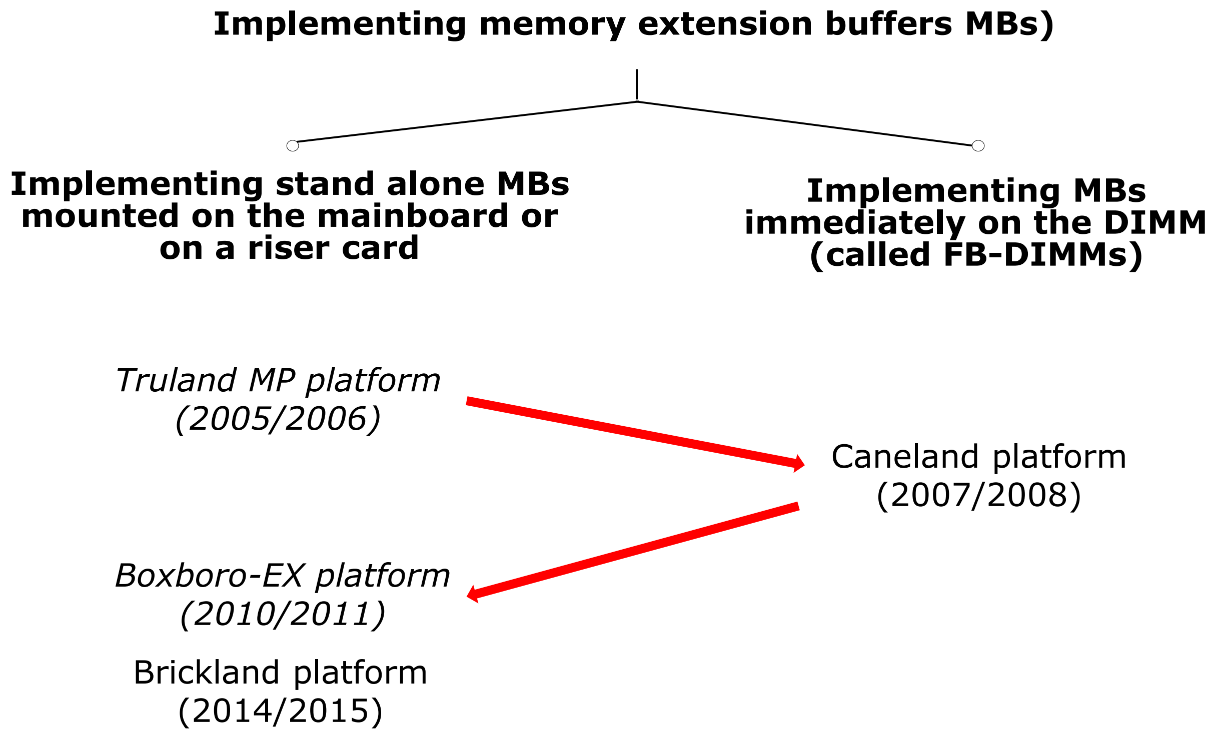


Figure: Intel's implementation of memory extension buffers

*

c) Evolution to the up to 10-core Boxboro-EX 4S/8S high-end server platform (targeting the Nehalem-EX/Westmere-EX processors) -2

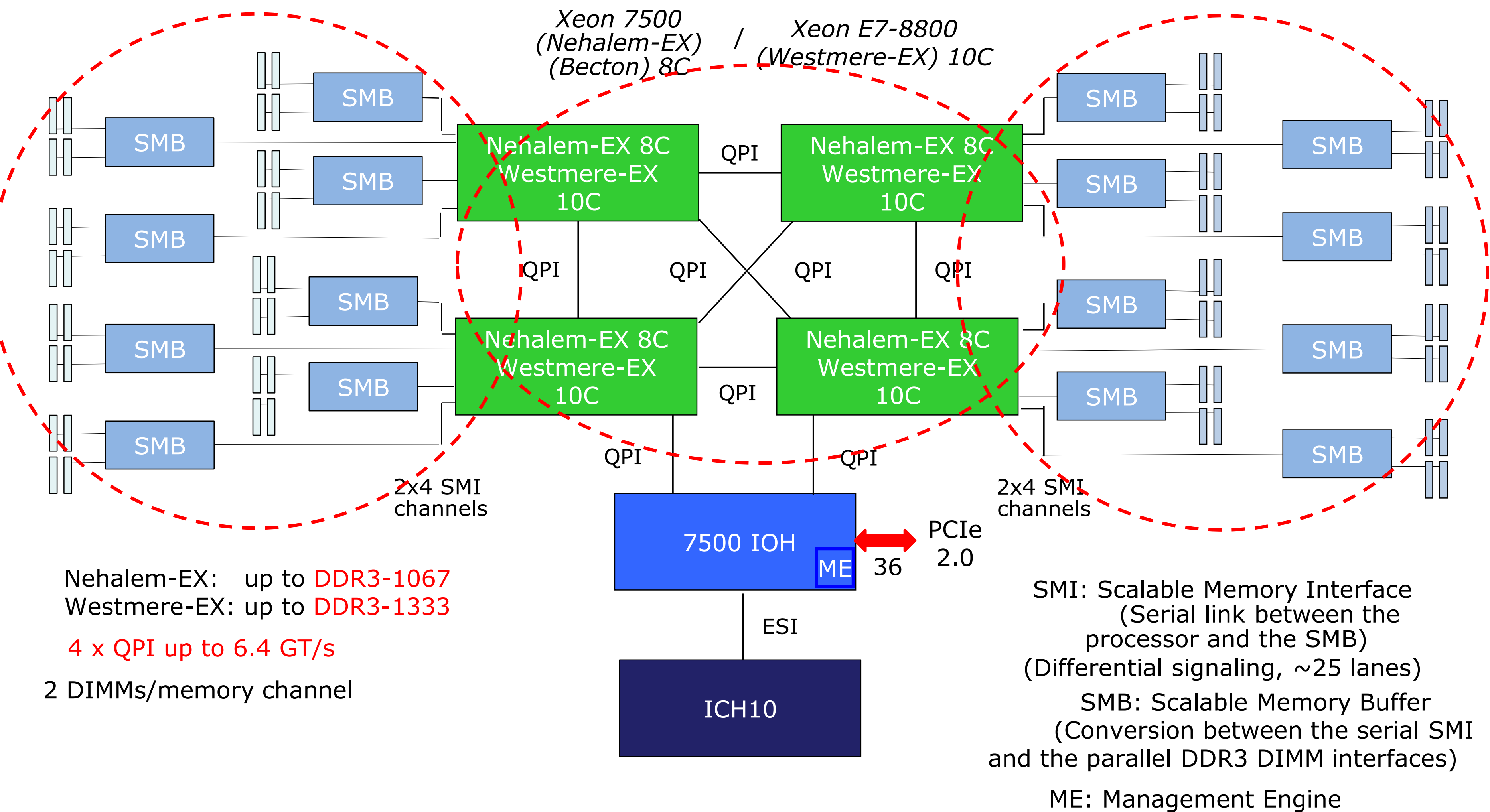


Figure: The Nehalem-EX, Westmere-EX aimed Boxboro-EX MP server platform (for up to 10 C) (2010/2011)

*

d) Evolution to the up to 24-core Brickland 4S/8S high-end server platform (targeting the Ivy Bridge-EX/Haswell-EX/Broadwell-EX processors) -1

Aims

a) Reducing power dissipation by eliminating serial/parallel conversions.

The previous Boxboro-EX platform made use of serial, packet based, differential links (called SMI links) between the memory controllers and the memory buffers that included about 70 signal lines.

In the subsequent Brickland platform Intel eliminated serial/parallel conversions and thus reduced power consumption by replacing serial links by proprietary parallel links, called SMI2 links.

Nevertheless, an interface conversion (between SMI 2 and DDR3/4) is needed further on.

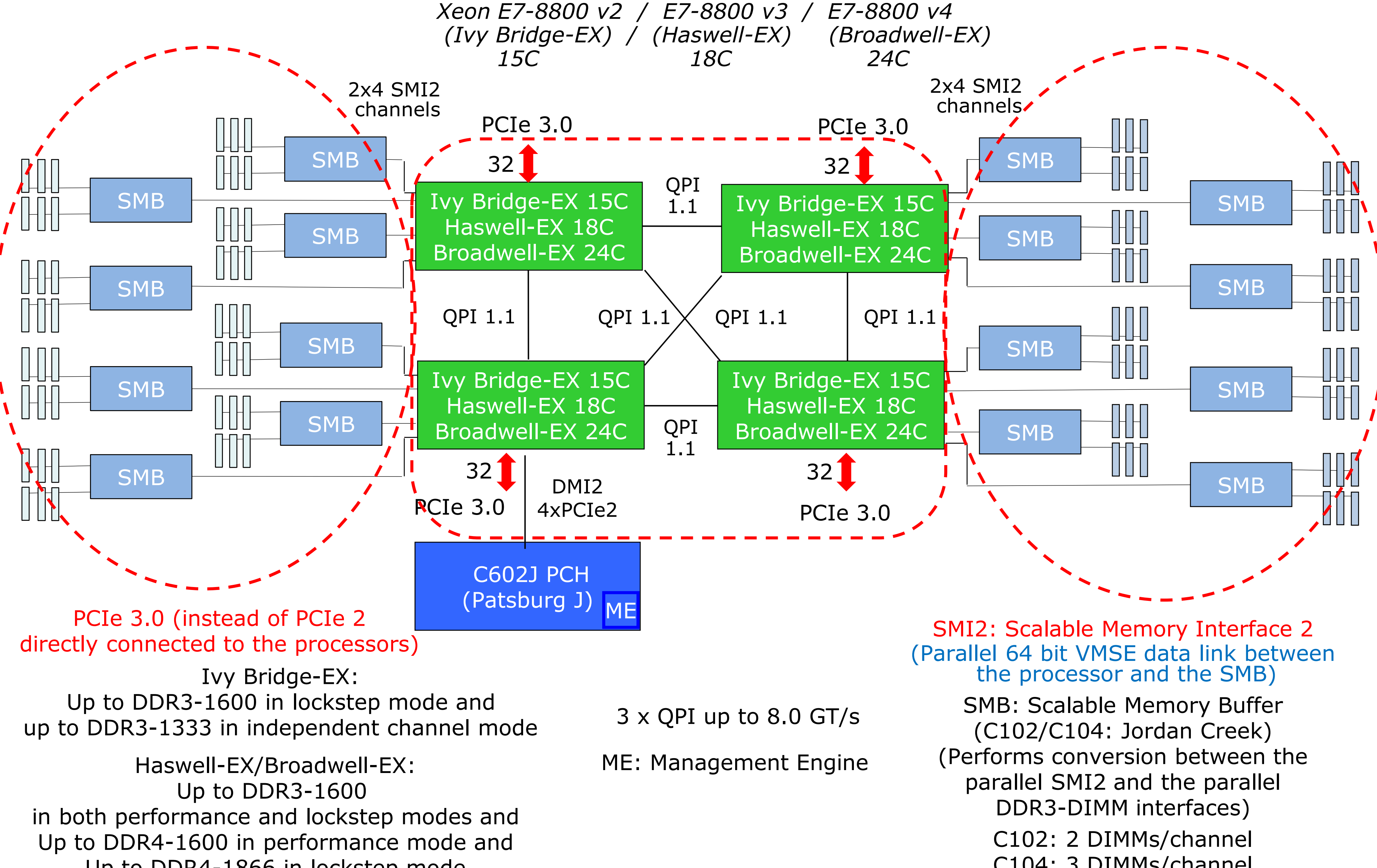
SMI 2 links provide 64-bit parallel, bidirectional communication using single-ended voltage reference signals, called VMSE (Voltage Mode Single Ended) signaling.

SMI 2 links require altogether about 110 signal lines, i.e. about 50 % more lines than SMI.

b) Providing faster, PCIe 3.0 lanes and connecting PCIe 3.0 lanes directly to the processors.

*

d) Evolution to the up to 24-core Brickland 4S/8S high-end server platform (targeting the Ivy Bridge-EX/Haswell-EX/Broadwell-EX processors) -2



Benefits of connecting PCIe 3.0 lanes directly to the processors

- Increasing the lane count from 36 to 4 x 32 = 128 lanes.
- Reducing the bandwidth demand between the processor and the chipset.

This allows

- to replace the high-bandwidth QPI link between the processor and the chipset by a low-bandwidth link (DMI2 link), comprising of 4 PCIe 2.0 lanes,
- to reduce the number of QPI links from 4 to 3 per processor and
- to integrate both chipset parts into a single PCH (Peripheral Control Hub).

*

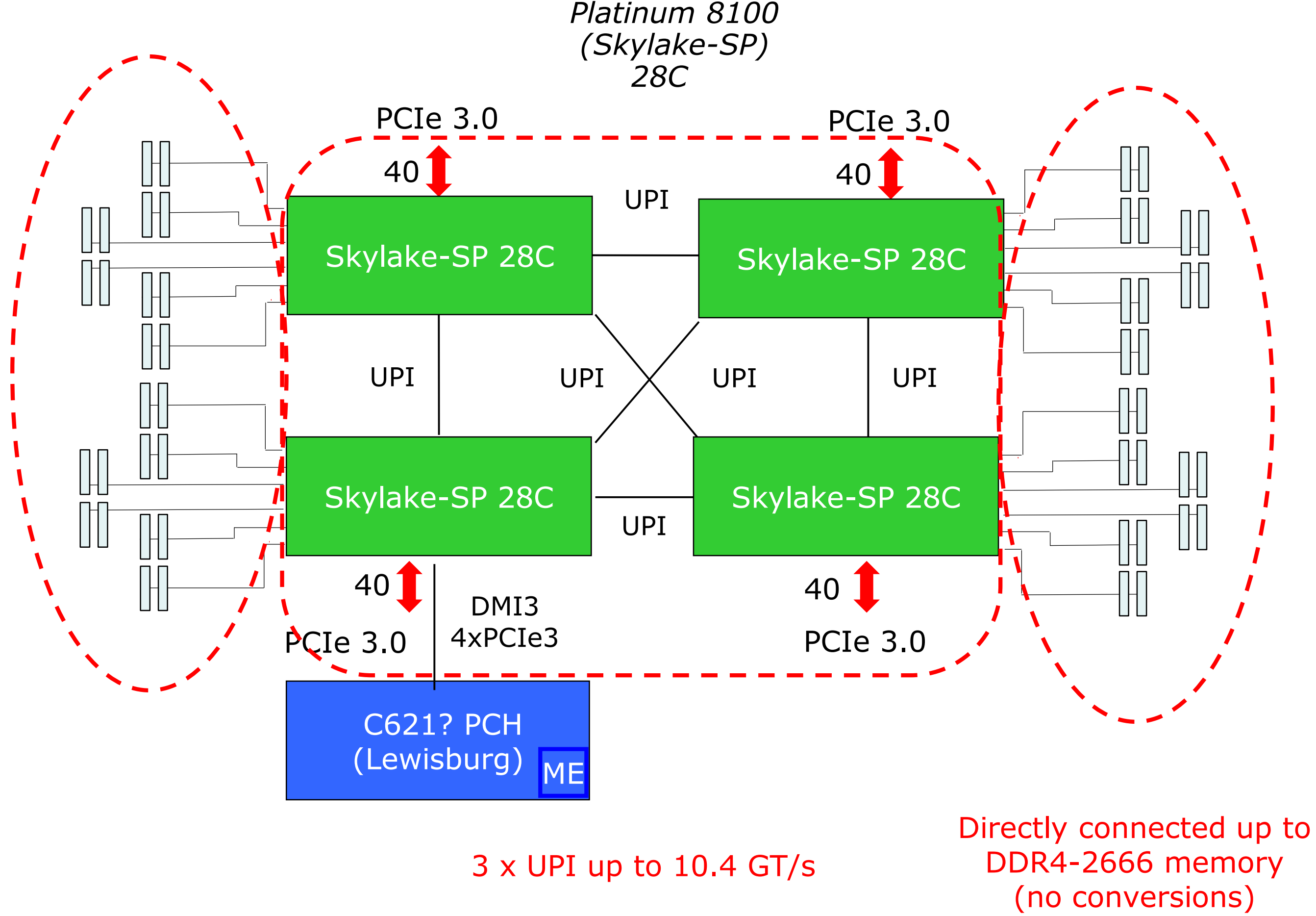
e) Evolution to the up to 28-core Purley high-end 2S/4S/8S server platform targeting the Skylake-SP processors) -1

Aim: Reducing power consumption and simplifying system implementation by eliminating interface conversions and related memory buffers by connecting DDR4 memory directly to the processors.

This requires a large socket (76x56 mm) with a high number of contacts (3647).

*

e) Evolution to the up to 28-core Purley high-end 2S/4S/8S server platform targeting the Skylake-SP processors) -2



3 x UPI up to 10.4 GT/s

Directly connected up to DDR4-2666 memory (no conversions)

*

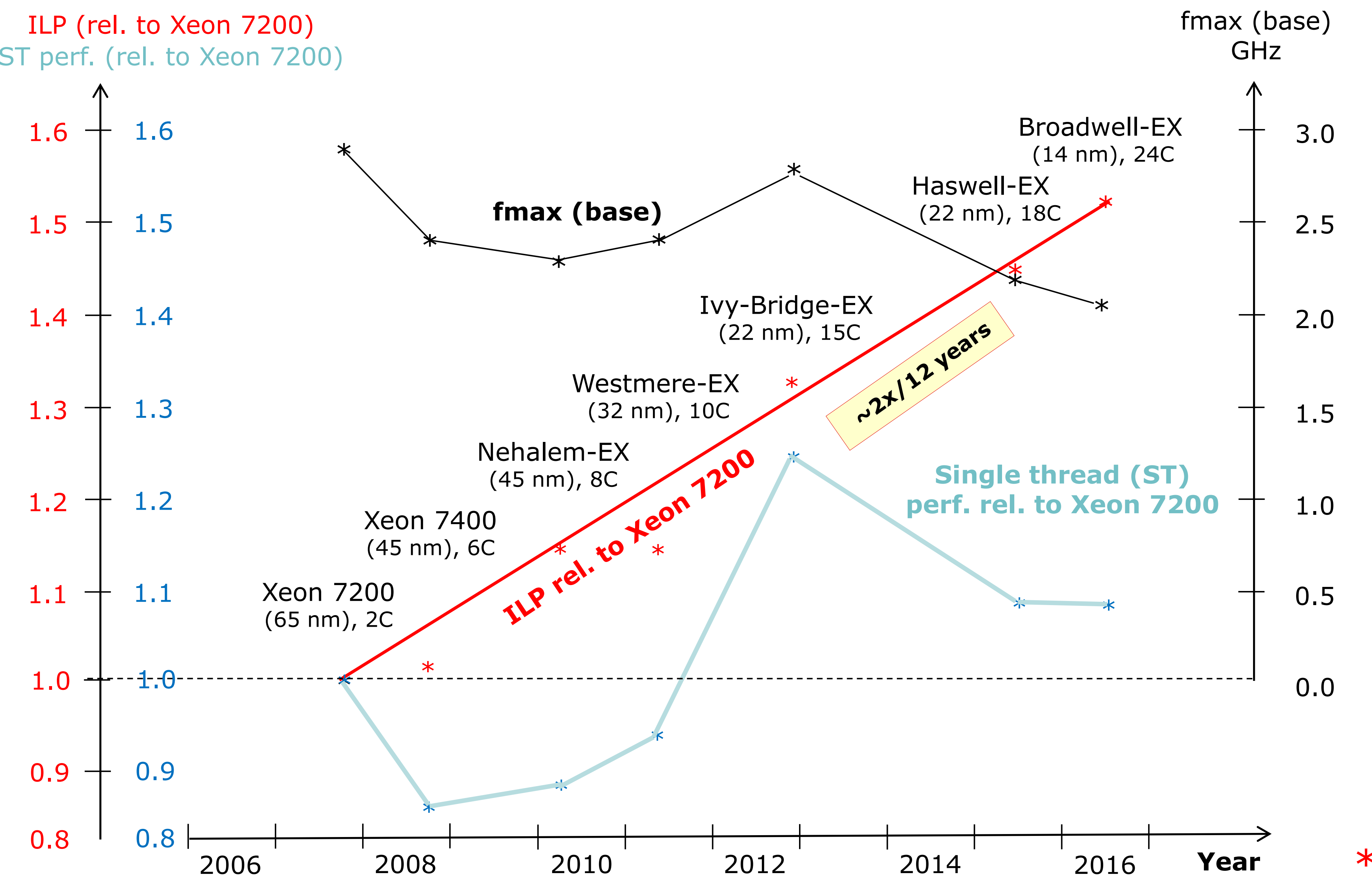
Intel's high-end 4S/8S server platforms – Cache architectures

Platform	4S platform topology	Date of intro.	Processor	Techno-logy	Core count (up to)	L2 cache	L3 cache (up to)
Truland MP	SMP w/dual FSBs	3/2005	90 nm Pentium 4 MP (Potomac)	90 nm	1C	1 MB	8 MB
		11/2005	Xeon 7000 (Paxville MP)	90 nm	2x1C	2 MB/C	--
		8/2006	Xeon 7100 (Tulsa)	65 nm	2C	1 MB /C	16 MB
Caneland	SMP w/Quad FSBs	9/2007	Xeon 7200 Tigerton DC (Core 2)	65 nm	2C	4 MB/C	--
		9/2007	Xeon 7300 Tigerton QC (Core 2)	65 nm	2x2C	4 MB/C	---
		9/2008	Xeon 7400 (Dunnington) (Penryn)	45 nm	6C	3 MB/C	16 MB
Boxboro-EX	NUMA fully connect. by 3 QPI buses	3/2010	Nehalem-EX (Xeon 7500/ (Beckton)	45 nm	8C	¼ MB/C	8x3 MB
		4/2011	Westmere-EX (E7-8800)	32 nm	10C	¼ MB/C	10x3 MB
Brickland	NUMA fully connect. by 3 QPI buses	2/2014	Ivy Bridge-EX (E7-8800 v2)	22 nm	15C	¼ MB/C	15x2.5 MB
		5/2015	Haswell-EX (E7-8800 v3)	22 nm	18C	¼ MB/C	18x2.5 MB
		6/2016	Broadwell-EX (E7-8800 v4)	14 nm	24C	¼ MB/C	24x2.5 MB
Purley	NUMA fully connect. by 3 UPI buses	7/2017	Skylake-SP (Platinum 8100)	14 nm	28C	1 MB/C.	28x1.375 MB

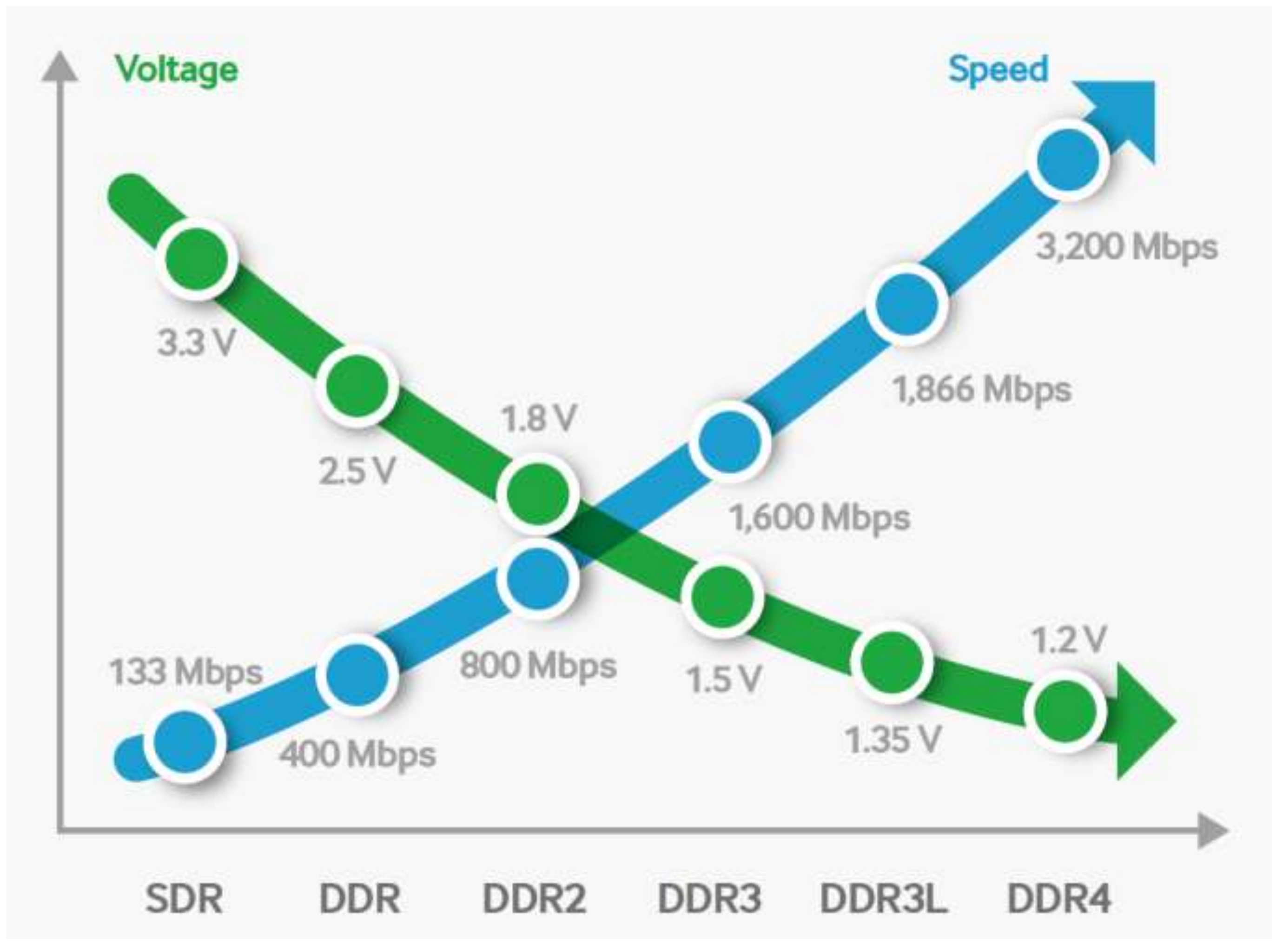
Intel's high-end 4S/8S server platforms – Connectivity

Platform	4S platform topology	Date of intro.	Processor	Techno-logy	Core count (up to)	FSB/QPI	IF to chipset	PCIe
Truland MP	SMP w/dual FSBs	3/2005	90 nm Pentium 4 MP (Potomac)	90 nm	1C	2xFSB 667 MT/s	HI 1.5 (266 MB/s bidirectional)	28 x PCIe lanes on the MCH
		11/2005	Xeon 7000 (Paxville MP)	90 nm	2x1C	2xFSB 800 MT/s		
		8/2006	Xeon 7100 (Tulsa)	65 nm	2C			
Caneland	SMP w/Quad FSBs	9/2007	Xeon 7200 Tigerton DC (Core 2)	65 nm	2C	4xFSB 1066 MT/s	ESI x4 (4x PCIe lanes) (1GB/s per direction)	28 x PCIe lanes on the MCH
		9/2007	Xeon 7300 Tigerton QC (Core 2)	65 nm	2x2C			
		9/2008	Xeon 7400 (Dunnington) (Penryn)	45 nm	6C			
Boxboro-EX	NUMA fully connect. by 3 QPI buses	3/2010	Nehalem-EX (Xeon 7500/ (Beckton)	45 nm	8C	4xQPI 6.4 GT/s (3 for NUMA)	QPI (6.4 GB/s)	36 x PCIe 2.0 lanes on the IOH
		4/2011	Westmere-EX (E7-8800)	32 nm	10C			
Brickland	NUMA fully connect. by 3 QPI buses	2/2014	Ivy Bridge-EX (E7-8800 v2)	22 nm	15C	3xQPI 1.1 8 GT/s	DMI2 x4 (4xPCIe 2.0) (2 GB/s per direction)	32 x PCIe 3.0 lanes on the proc.
		5/2015	Haswell-EX (E7-8800 v3)	22 nm	18C	3xQPI 1.1 9.6 GT/s		
		6/2016	Broadwell-EX (E7-8800 v4)	22 nm	24C	3xQPI 1.1 9.6 GT/s		
Purley	NUMA fully connect. by 3 UPI buses	2017	Skylake-EX	14 nm	28C	3xUPI 10.4 GT/s	DMI3 x4 (4xPCIe 3.0) (~ 4GB/s/dir.	48 x PCIe lanes on the proc.

Intel's high-end 4S/8S server platforms – Performance features (Source of data: Intel)



Supply voltage and max. transfer rate (speed) of major DRAM types [136]



*

3. Example 1: The Brickland platform

- 3.1 Overview of the Brickland platform
- 3.2 Key innovations of the Brickland platform vs. the previous Boxboro-EX platform
- 3.3 The Ivy Bridge-EX (E7-8800 v2) 8S processor line
- 3.4 The Haswell-EX (E7-8800 v3) 8S processor line
- 3.5 The Broadwell-EX (E7-8800 v4) 8S processor line

It won't be discussed.

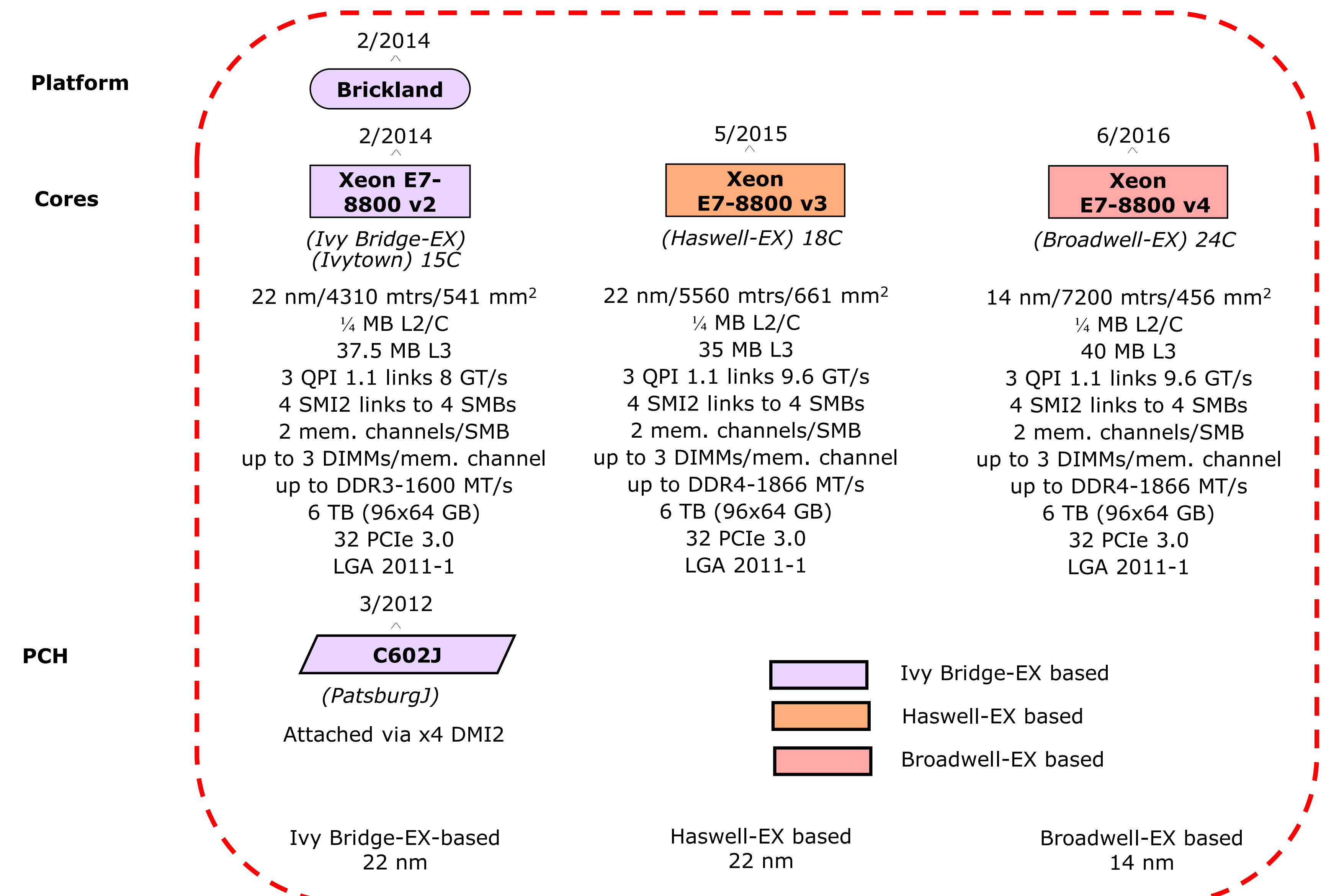
3.1 Overview of the Brickland platform

3.1 Overview of the Brickland platform

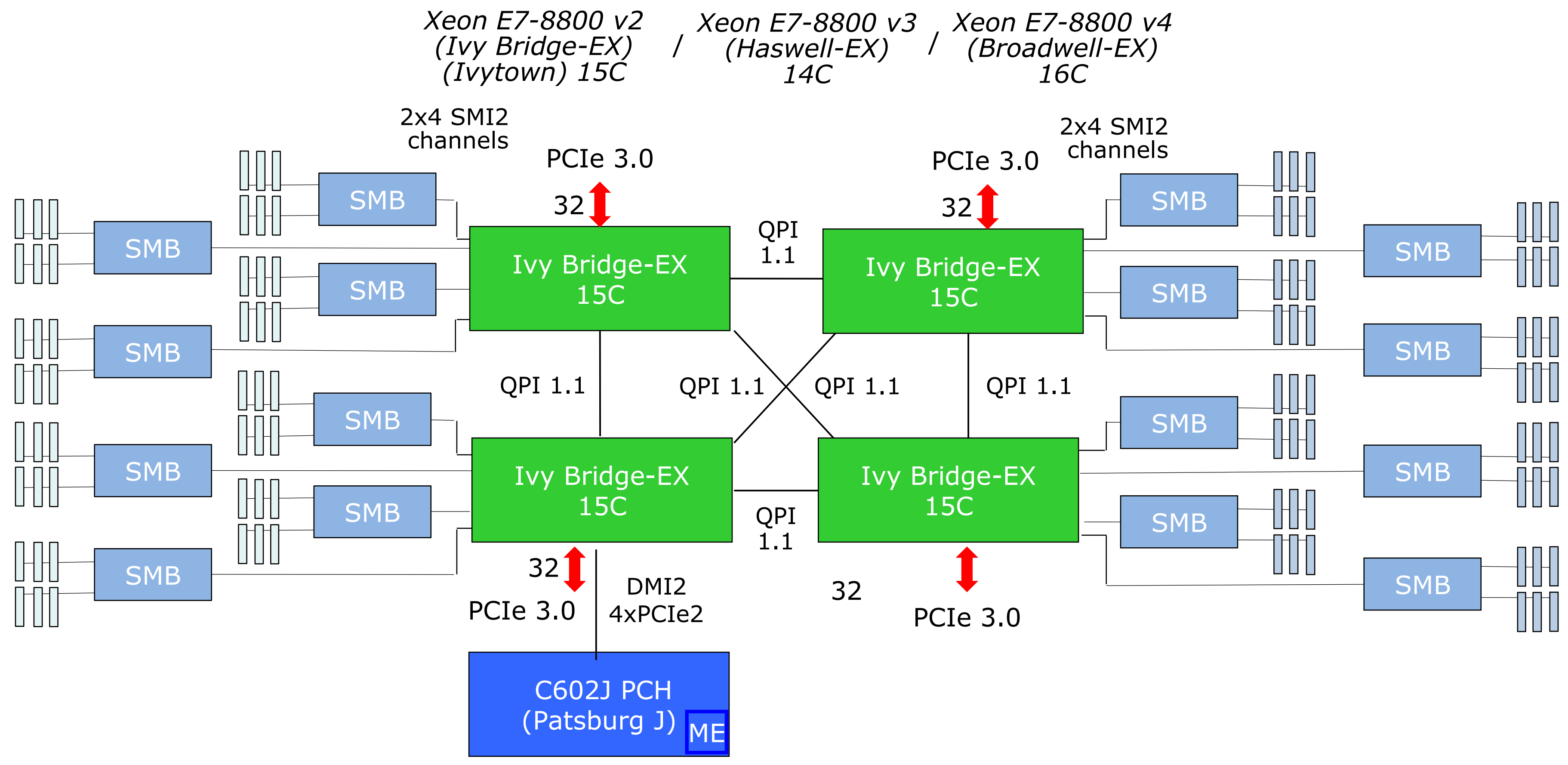
Overview of Intel's high-end 4S/8S server platforms and processor lines

Platform	Core	Techn.	Intro.	high-end 4S/8S server processor lines	Core count	Chipset	Proc. socket
Truland MP	Pentium 4 MP Prescott	90 nm	3/2005	90 nm Pentium 4 MP (Potomac)	1C	E8500 + ICH5	LGA 604
	Pentium 4 Presc.	90 nm	11/2005	7000 (Paxville MP)	2x1C	E8501 + ICH5	
	Pentium 4 Presc.	65 nm	8/2006	7100 (Tulsa)	2x1C		
Caneland MP	Core2	65 nm	9/2007	7200 (Tigerton DC) 7300 (Tigerton QC)	2C 2x2C	E7300 (Clarksboro)+ 631x/632x ESB	LGA 604
	Penryn	45 nm	9/2008	7400 (Dunnington)	6C		
Boxboro-EX	Nehalem	45 nm	3/2010	7500 (Beckton/ Nehalem-EX)	8C	7500 (Boxboro) + ICH10	LGA 1567
	Westmere	32 nm	4/2011	E7-8800 (Westmere-EX)	10C		
	Sandy Bidge	32 nm					
Brickland	Ivy Bridge	22 nm	2/2014	E7-8800 v2 (Ivy Bridge-EX)	15C	C602J (Patsburg J)	LGA 2011-1
	Haswell	22 nm	5/2015	E7-8800 v3 (Haswell-EX)	18C		
	Broadwell	14 nm	6/2016	E7-8800 v4 (Broadwell-EX)	24C		
Purley	Skylake	14 nm	7/2017	Platinum 8100 (Skylake-SP)	28C	C620 (Lewisburg)	LGA 3647

The Brickland (8S) platform



Basic system architecture of the high-end 4S/8S Brickland server platform
(assuming 2x3 DIMMs/SMB)



SMI2: Scalable Memory Interface 2
(Parallel 64 bit VMSE data link between the processor and the SMB)

SMB: Scalable Memory Buffer
(Performs conversion between the parallel SMI2 and the parallel DIMM interfaces)

ME: Management Engine (Dedicated microprocessor to manage the server via a console or remotely)

3.2 Key innovations of the Brickland platform vs. the previous Boxboro-EX platform

3.2 Key innovations of the Brickland platform vs. the previous Boxboro-EX platform

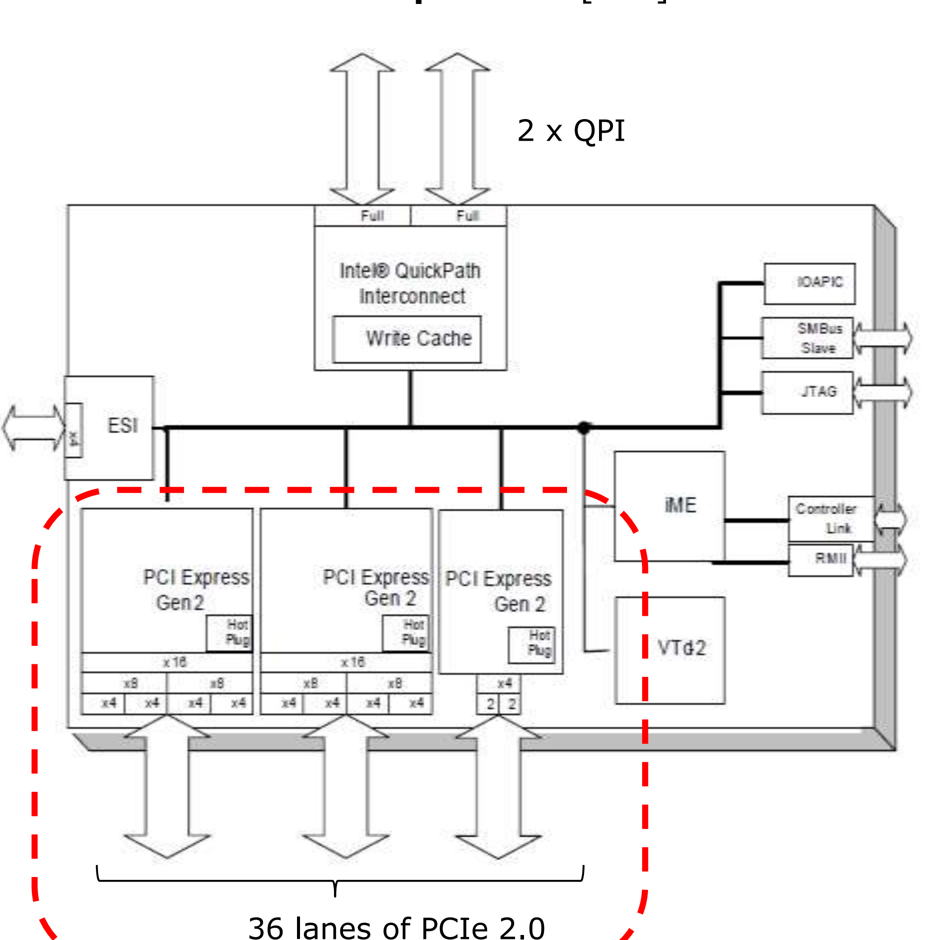
3.2.1 Connecting PCIe links direct to the processors rather than to the MCH

3.2.2 New memory buffer design

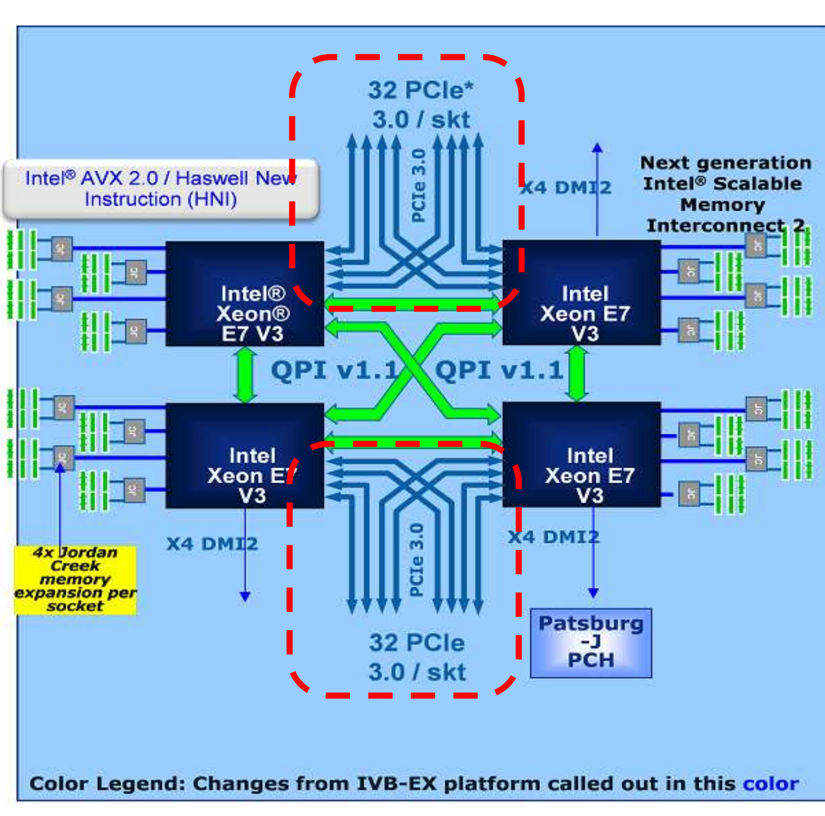
3.2.1 Connecting PCIe links direct to the processors rather than to the MCH -1

3.2.1.1 Overview

Boxboro platform [106]



Brickland platform [114]



Connecting PCIe links direct to the processors rather than to the MCH -2

It has two major implications, including

- Raising the PCIe lane count provided by the Brickland platform (see Section 3.2.1.2)
- Reducing the bandwidth needed between the processors and the chipset (see Section 3.2.1.3)

3.2.1.2 Raising the PCIe lane count provided by the Brickland platform -1

As long as the Boxboro-EX platform provides 36 PCIe lanes, the Brickland platform affords 32 PCIe lanes per processor, i.e. 4 x 32 = 132 lanes in total for a 4S platform, as shown below.

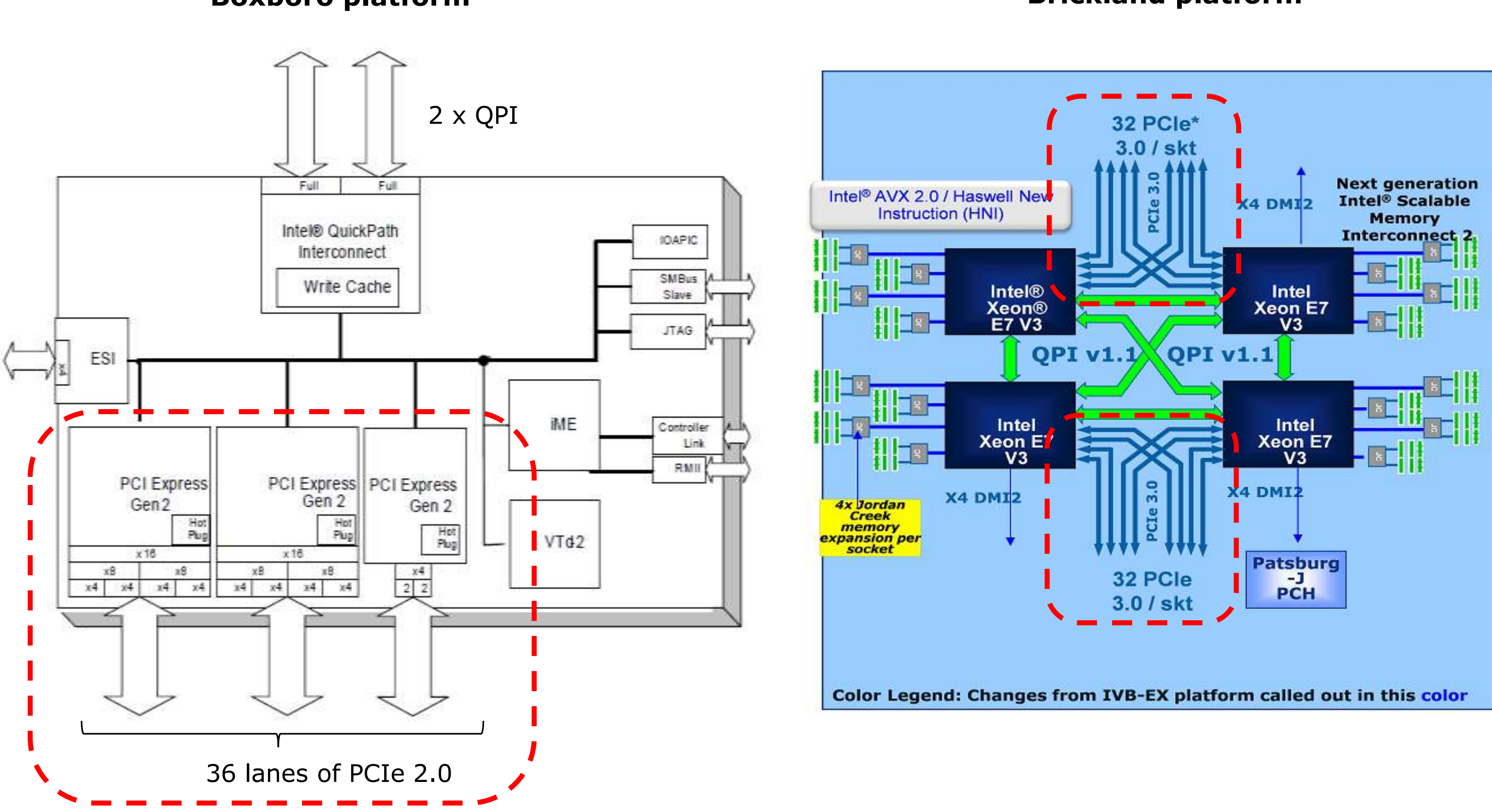


Figure: Connecting PCIe links direct to the processors rather than to the MCH

3.2.1.2 Raising the PCIe lane count provided by the Brickland platform -2

- If the PCI lanes are connected directly to the processors rather than to the MCH, the number of PCIe lanes supported by the platform will scale linearly with the number of processors.
- In addition, the Brickland platform provides faster PCIe 3.0 lanes instead of PCI 2.0 lanes of the previous Boxboro platform.

  Per lane PCIe bandwidth:

  - PCI 2.0: 500 MB/s
  - PCI 3.0: 984.6 MB/s

3.2.1.3 Reducing the bandwidth demand between the processors and the chipset

- By relocating the PCIe links from the MCH to the processors, the bandwidth needed between the chipset and the processors becomes significantly reduced, as follows;

  In the preceding Boxboro platform the MCH provides 36 PCIe 2.0 lanes with a total bandwidth of 36x0.5 GB/s = 18 GB/s per direction.

  This gives rise for using a QPI 1.1 bus width the bandwidth of 16.0 GB/s per direction between the MCH and the processors.

- By contrast, in the Brickland platform the PCIe lanes are connected immediately to the processors and there is no need for providing the associated bandwidth.
- This has three consequences;

  a) Interconnecting the processors and the PCH by a DMI2 (4x PCIe 2.0) link

  b) Providing only three QPI links per processor

  c) Implementing a single chip “chipset”,

  as discussed next.

a) Interconnecting the processors and the PCH by a DMI2 (4x PCIe 2.0) interface

Due to the reduced bandwidth demand of the chipset-processor interface, it suffices now to interconnect the PCH with a single processor via a DMI2 interface consisting of 4 PCIe 2.0 lanes that provide a bandwidth of up to 4x0.5=2 GB/s, rather than using a QPI bus with a bandwidth of 16.0 GB/s.

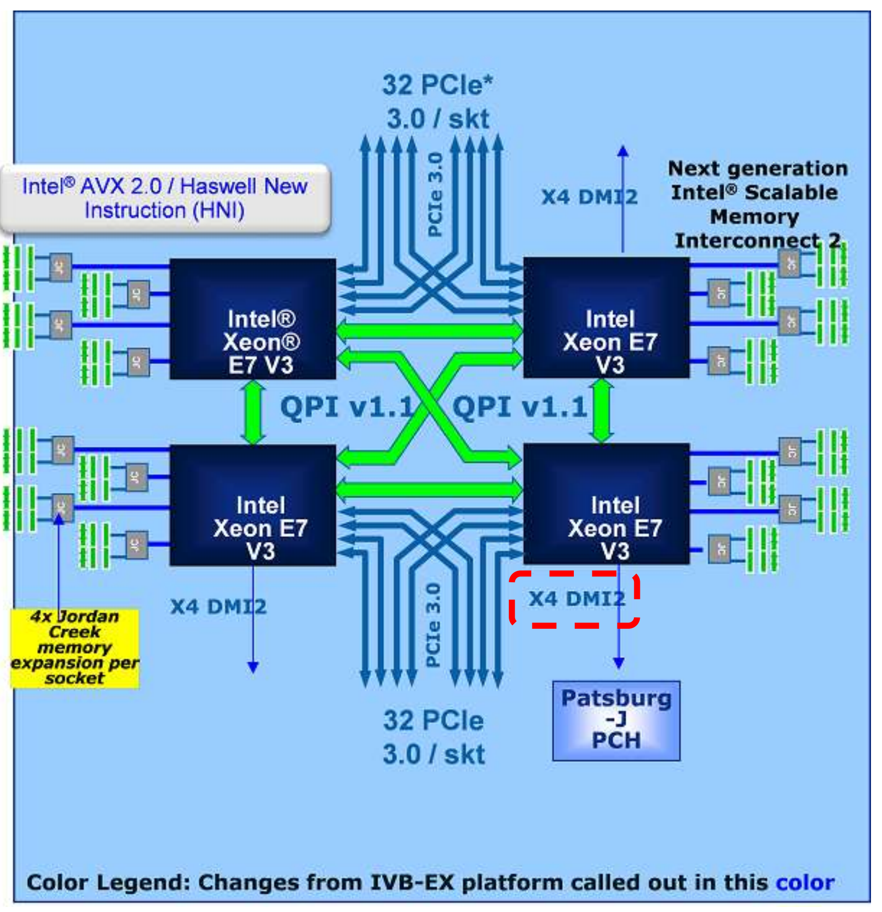


Figure: The Haswell-EX implementation of the Brickland platform [114]

b) Providing only three QPI links per processor -1

As the Brickland platform does not need an extra QPI bus to interconnect the processor with the PCH, it has only 3 QPI buses per processor rather than four compared to the previous (Boxboro-EX) platform, as shown in the next Figure.

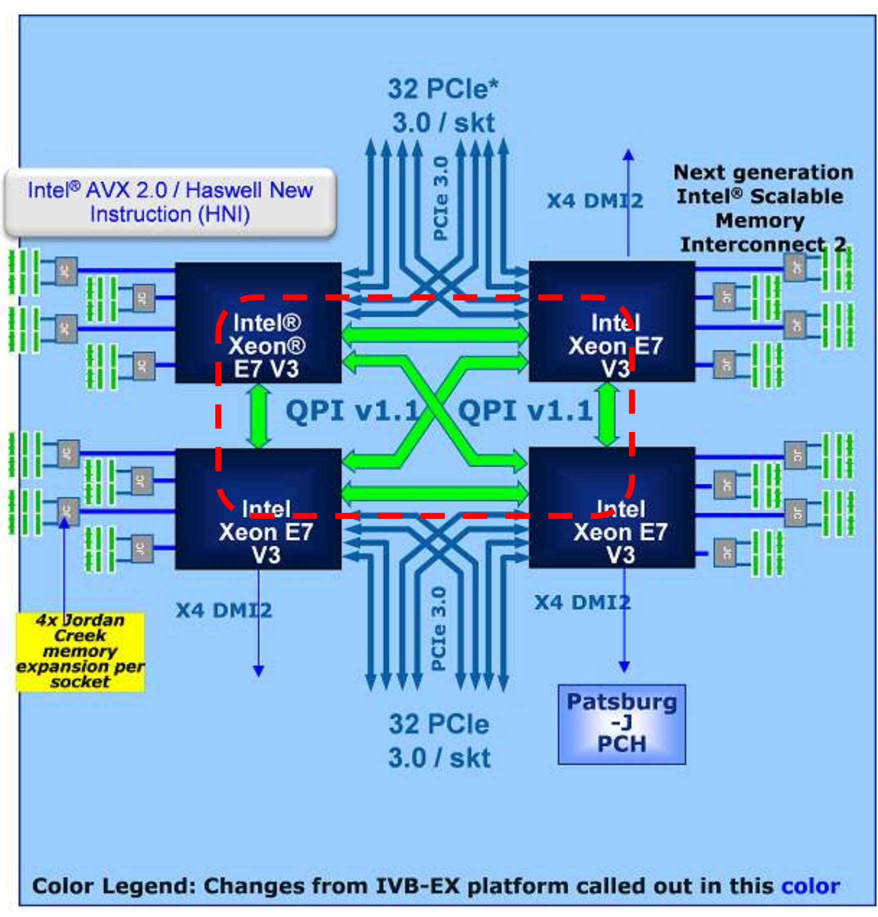


Figure: The Haswell-EX implementation of the Brickland platform [114]

3.2.1 Connecting PCIe links direct to the processors rather than to the MCH (8)

Providing only three QPI links per processor -2

In addition, the Brickland platform supports already QPI 1.1 buses with the following speeds:

- Ivy Bridge-EX based Brickland platforms: up to 8.0 Gbps
- Haswell-EX based Brickland platforms: up to 9.6 Gbps

c) Implementing a single chip "chipset"

- Inspired by the reduced functionality, the chipset is now implemented as a single chip solution instead of using two chips, as in the previous Boxboro platform.
- The single chip "chipset" is now designated as the PCH (Platform Controller Hub). It is in particular the C602 J or Patsburg-J PCH, as shown below.

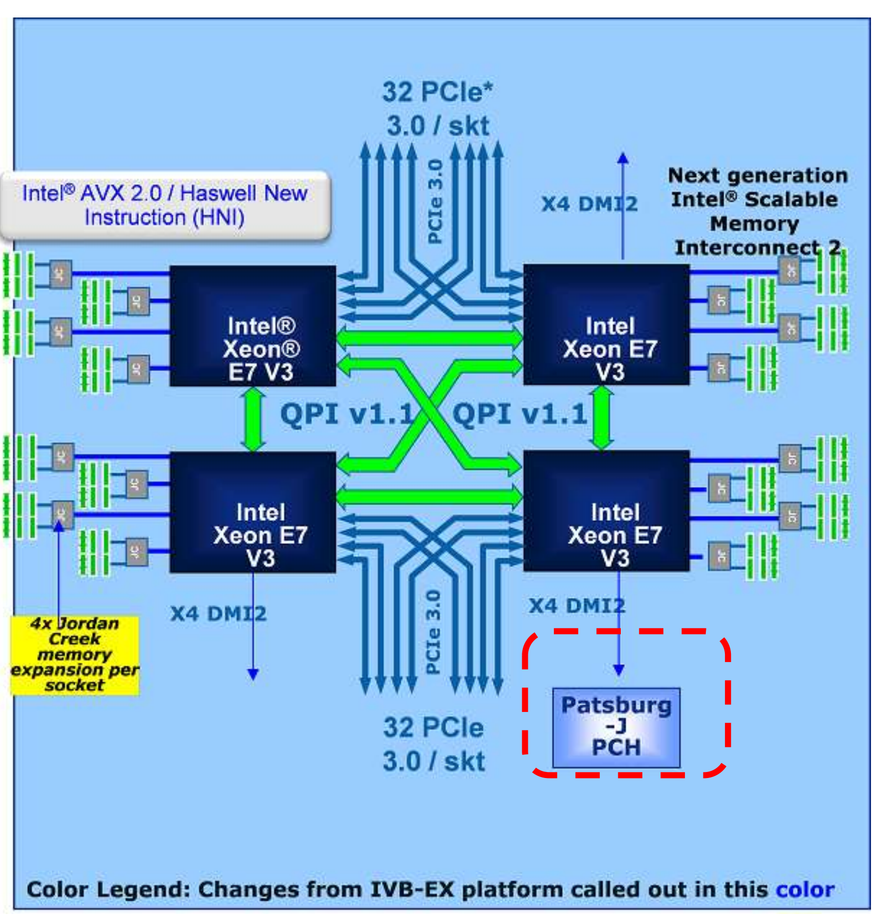


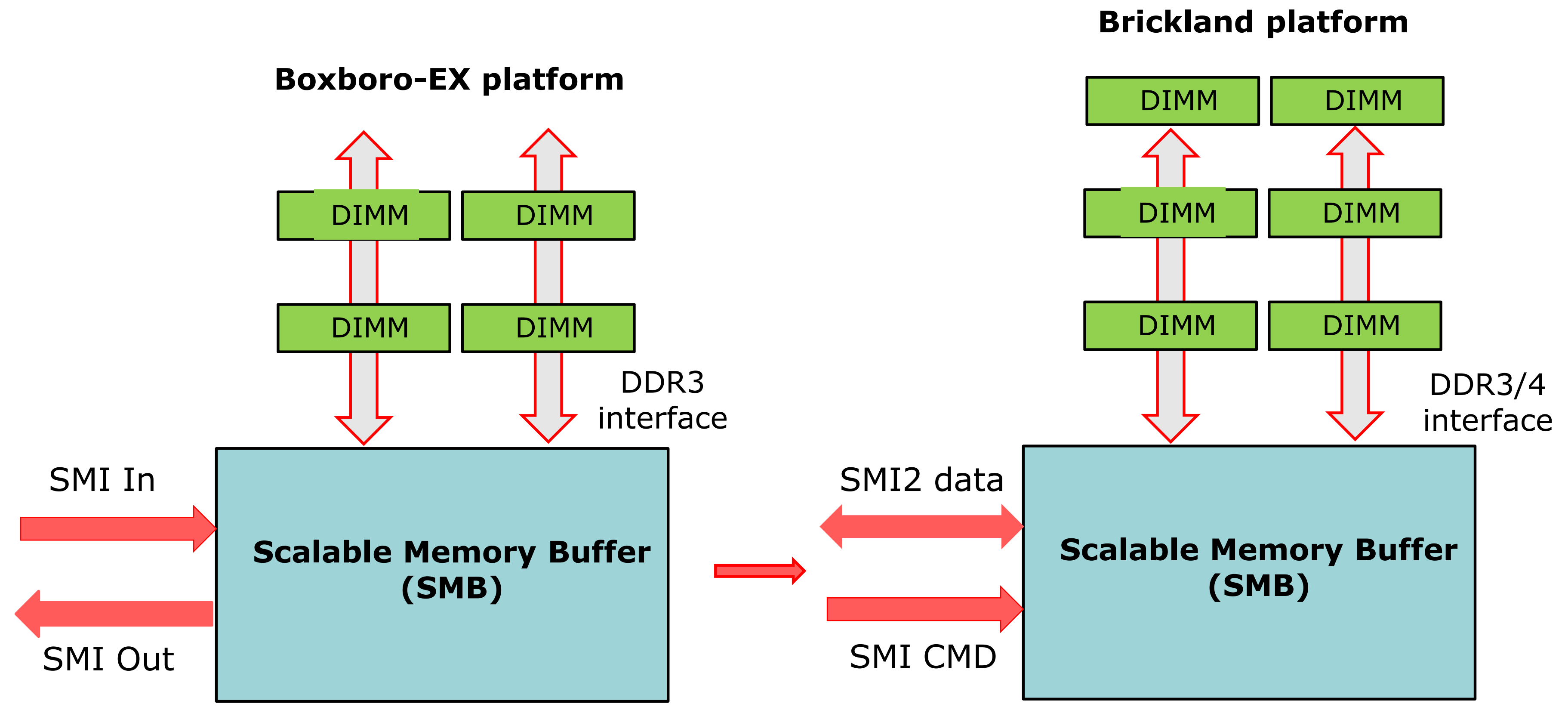
Figure: The Haswell-EX implementation of the Brickland platform [114]

3.2.2 New memory buffer design

It has two main components, as follows:

a) Redesigned, basically parallel MC-SMB interface, called SMI2 interface
b) Enhanced DRAM interface

a) Redesigned, basically parallel MC-SMB interface, called SMI2 interface -1



SMI: Serial, packet based link with differential signaling

Up to 6.4 Gbps

SMI 2: 64-bit parallel, bidirectional data link with single-ended voltage reference signaling, (VMSE (Voltage Mode single Ended) signaling)

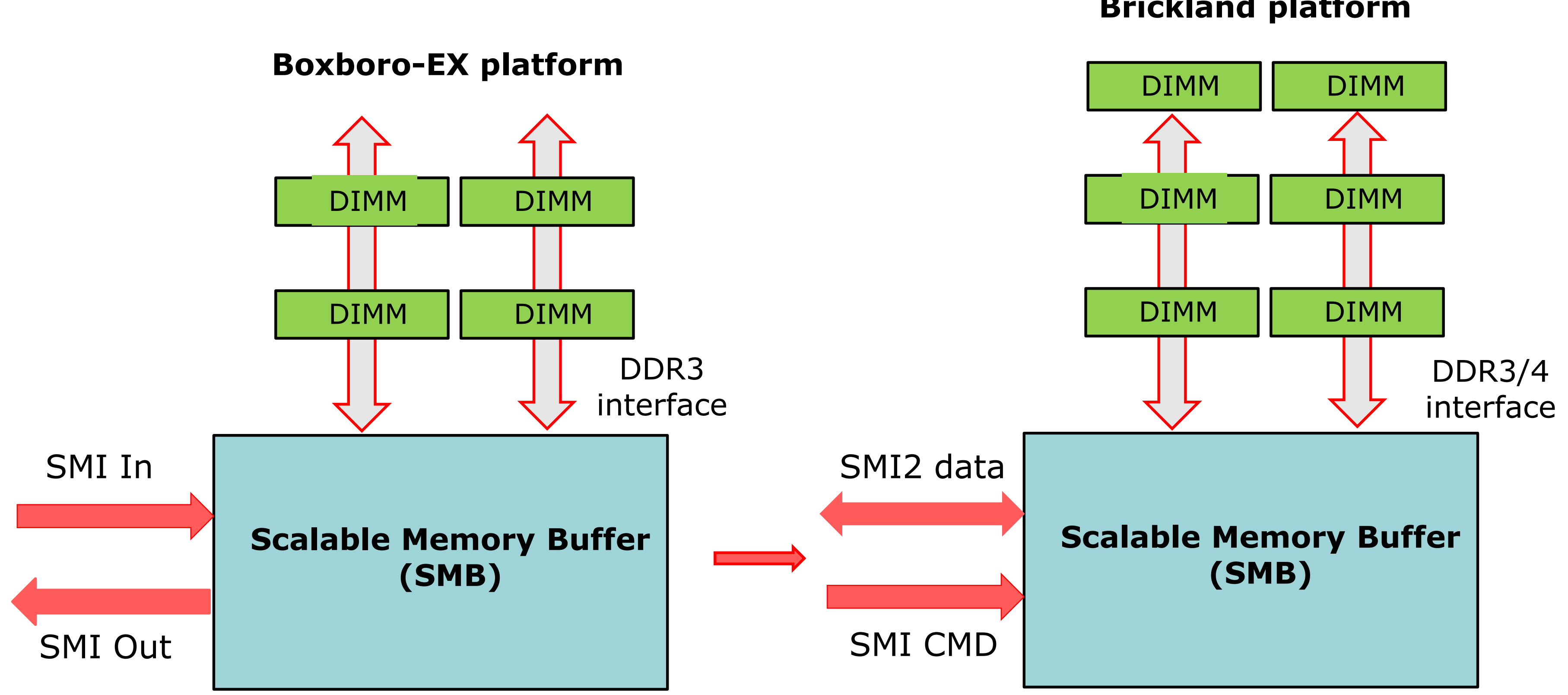
Up to 3200 MT/s

3.2.2 New memory buffer design (3)

a) Redesigned, basically parallel MC-SMB interface, called SMI2 interface -2

- The SMI link of the Boxboro-EX platform was a serial, packet based, differential link including about 70 signal lines.
- By contrast, as far as the data transfer is concerned, with the SMI2 link Intel changed to 64-bit parallel, bidirectional communication using single-ended voltage reference signals, called VMSE (Voltage Mode Single Ended) signaling.
- SMI 2 requires altogether about 110 signal lines, that is about 50 % more lines than SMI.
- The reason for changing from serial transfer to parallel transfer in case of the data lines is presumably the fact, that in this case AD/DA converting of data becomes superfluous and this results in reduced dissipation.
- The SMI 2 interface is used in the Brickland platform and operates
  - in Ivy Town-EX based platforms at a speed of up to 2667 MT/s,
  - in Haswell-EX based platforms at a speed of up to 3200 MT/s and
  - in Broadwell-EX based platforms at a speed of up to 3200 MT/s.

b) Enhanced DRAM interface



- Up to DDR3-1333
- Up to 2 DIMMs per memory channel

- Up to DDR4-1866
- Up to 3 DIMMs per memory channel

3.2.2 New memory buffer design (5)

Operation modes of the SMBs in the Brickland platform [113]

Operation modes of the SMBs

Lockstep mode

- In lockstep mode the same command is sent on both DRAM buses.
- The read or write commands are issued simultaneously to the referenced DIMMs, and the SMB interleaves the data on the

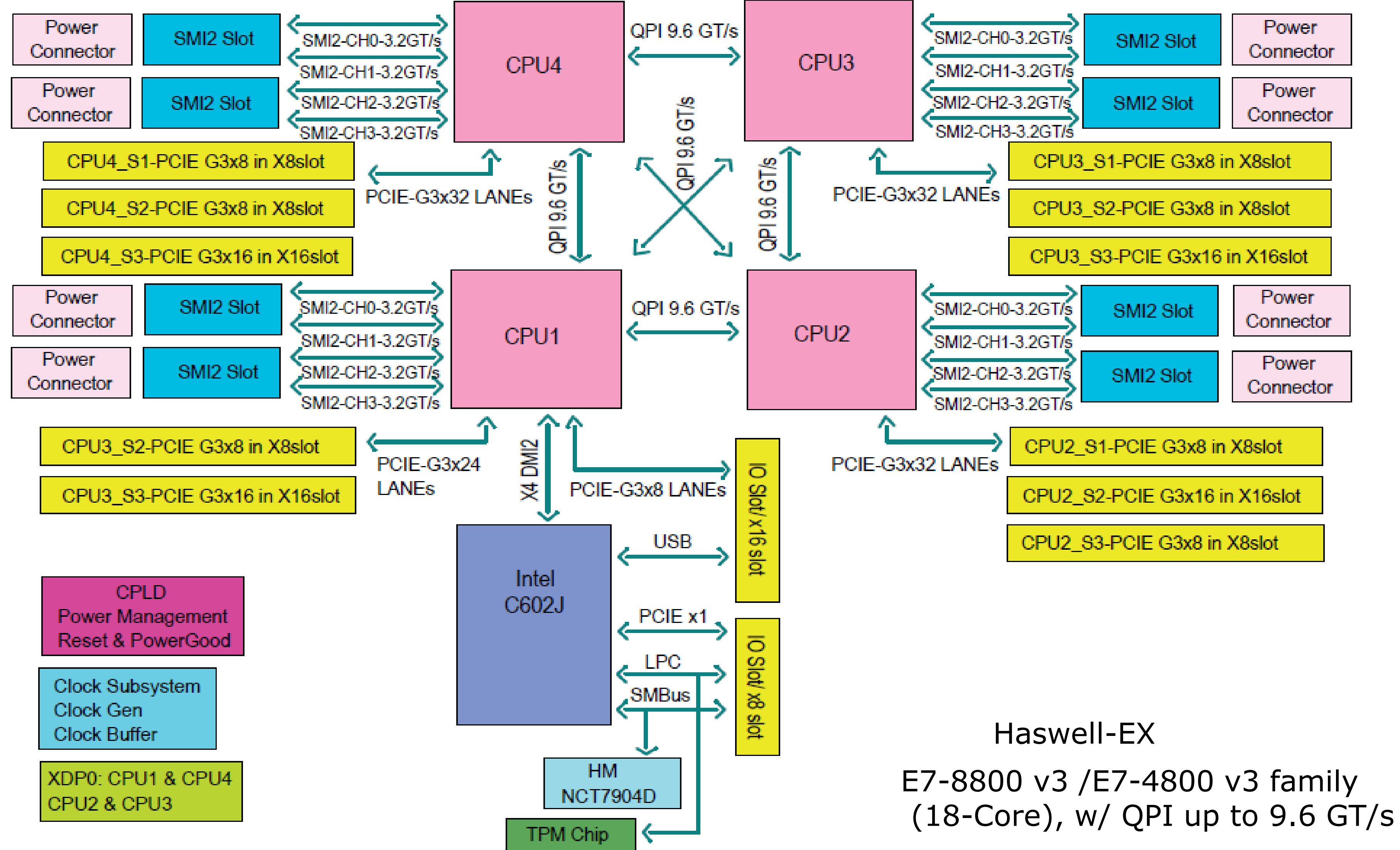
Independent channel mode (aka Performance mode)

- In the independent channel mode commands sent to both DRAM channels are independent from each other.

DRAM speeds of the Brickland platform

Feature	Ivy Bridge-EX based (E7-8800 v2)	Haswell-EX based (E7-8800 v3)
SMB	C102/C104	C112/C114
DDR4 speed	n.a.	Perf. mode: up to 1600 MT/s Lockstep mode: up to 1866 MT/s
DDR3 speed	Perf. mode: up to 1333 MT/s Lockstep mode: up to 1600 MT/s	Perf. mode: up to 1600 MT/s Lockstep mode: up to 1600 MT/s

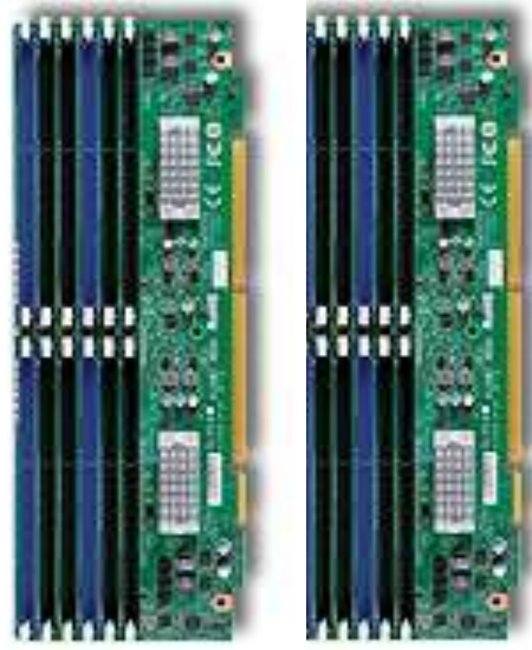
Example for a Haswell-EX based 4S Brickland platform [115]



The SMBs with the DIMMs are installed on extra cards to be inserted into the 8 SMI2 slots.

3.2.2 New memory buffer design (8)

The mainboard and two of the memory cards to be inserted into the SMI2 slots of the platform shown in the previous Figure [115]



Memory cards

Mainboard

3.3 The Ivy Bridge-EX (E7-8800 v2) 4S/8S processor line

3.3 The Ivy Bridge-EX (8800 v2) 4S/8S processor line [116]

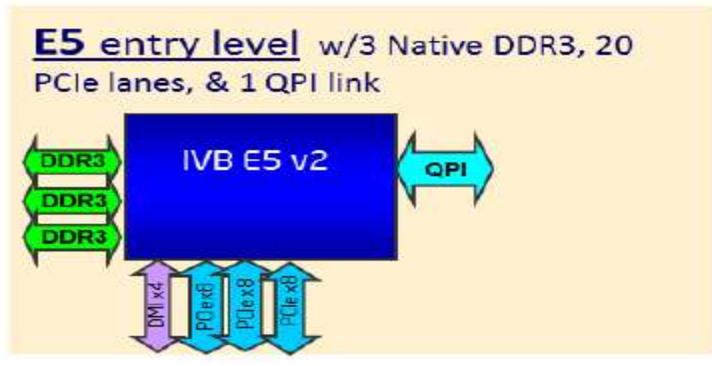
- Intel developed a single processor to cover all server market segments from 1S to 8S, called the Ivytown processor.
- It is manufactured on the 22 nm technology, includes up to 15 cores built up on 4.31 billion transistors on a die of 541 mm^2.
- Ivytown processor versions have a TDP of 40 to 150 W and operate at frequencies ranging from 1.4 to 3.8 GHz.
- The Ivytown processor was launched along with the Romley 2S server platform in 9/2013 followed by the Brickland 4S/8S platform in 2/2014.

3.3 The Ivy Bridge-EX (E7-8800 v2) 4S/8S processor line (2)

Server platforms and processor segments covered by the Ivytown processor [117]

E5 platform: Romley

Entry level
E5 2400 1, 2 sockets



Performance level
E5 2600 1, 2, 4 sockets

E5 w/4 Native DDR3, 40 PCIe lanes, & 2 QPI links

DDR3
DDR3
DDR3
DDR3
IVB E5 v2
QPI
QPI
DMI x4
PCIe x8
PCIe x8
PCIe x8
PCIe x8
PCIe x8

Socket R

E7 platform: Brickland

High-performance level
E7 4800/8800 2, 4, 8 sockets

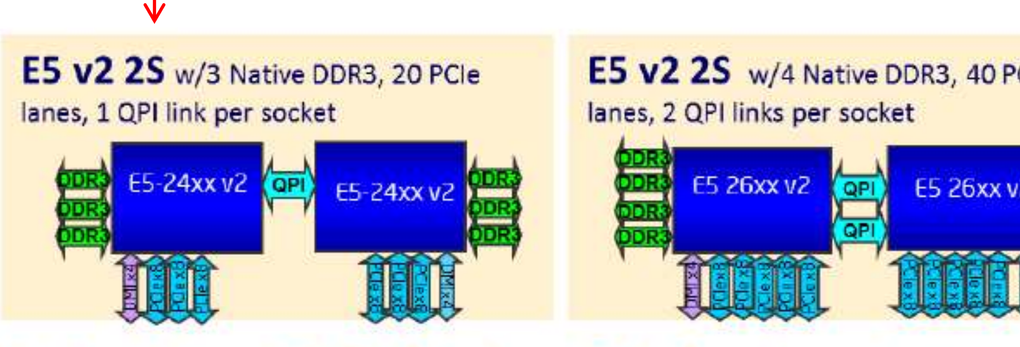
E7 w/8 Native DDR3 through Memory eXtension Buffer chips (MXB), 32 PCIe lanes, & 2 QPI links

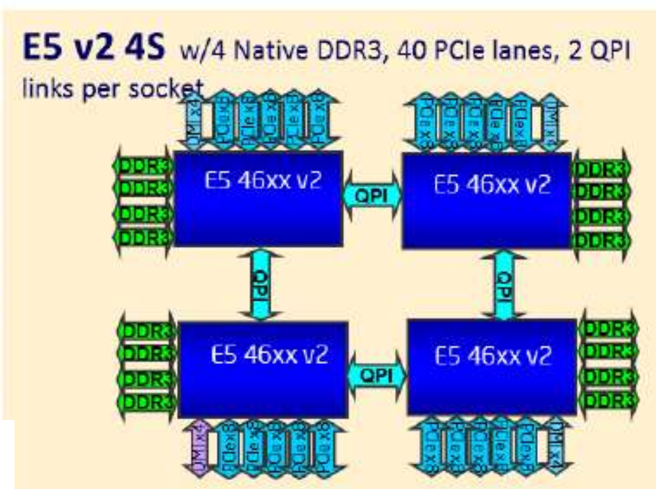
DDR3
DDR3
DDR3
DDR3
DDR3
DDR3
DDR3
DDR3
MXB
MXB
MXB
MXB
SMI2
SMI2
SMI2
SMI2
IVB E7 v2
QPI
QPI
QPI
DMI x4
PCIe x8
PCIe x8
PCIe x8
PCIe x8

Socket R2

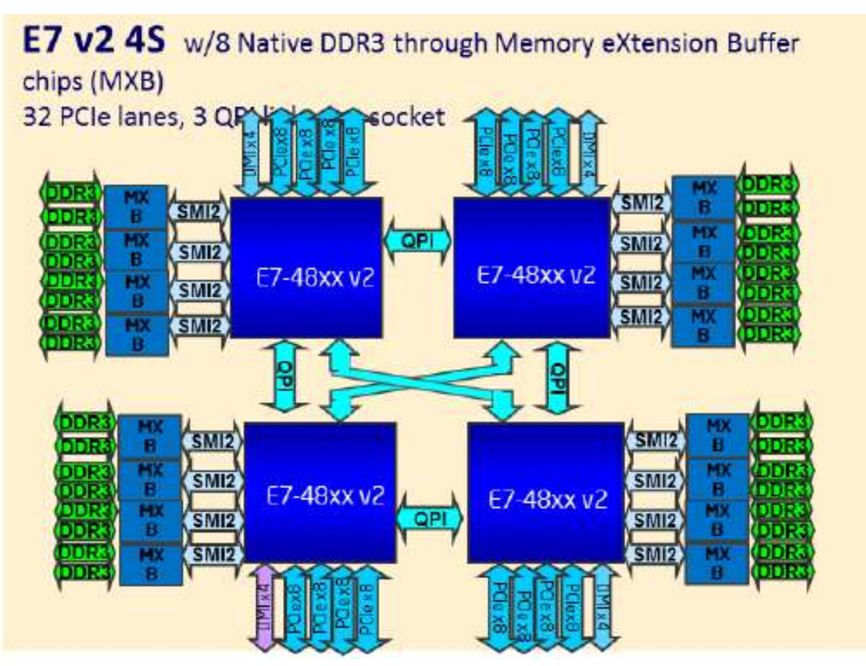
E5, E7 platform alternatives built up of Ivytown processors [117]

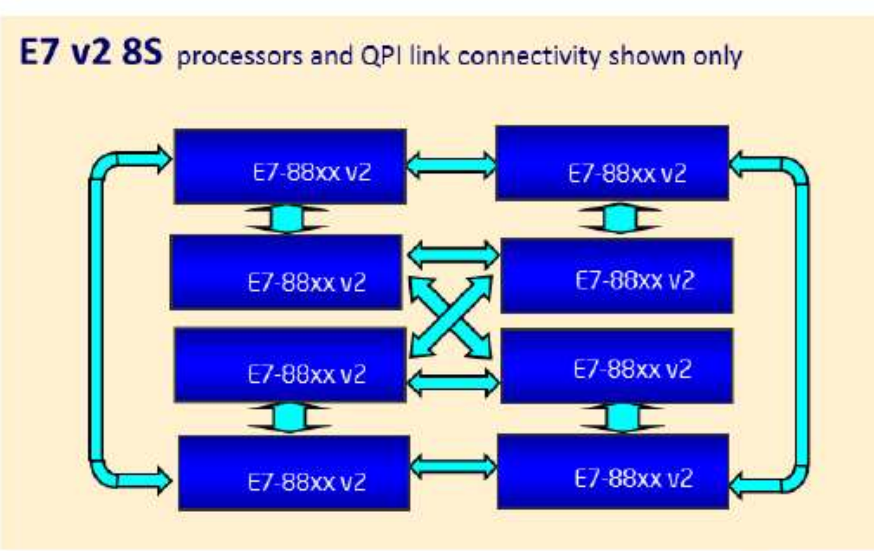
Romley Entry and Performance level platform alternatives





Brickland High performance platform alternatives





Block diagram of the Ivy Town processor

The processor consists of 3 slices, each with 5 cores and associated LLC segments, as seen below.

PCU: Power Control Unit
TAP: Test Access Port
FUSE: Fuse block, to configure the die i.e. to have various core and cache sizes as well as different frequency and TDP levels
DRNG: Digital Random Number Generator
IIO: Integrated I/O block
HA: Home Agent providing memory coherency
MC: Memory Controller
VMSE: Voltage-Mode Single-Ended Interface (actually the SMB)

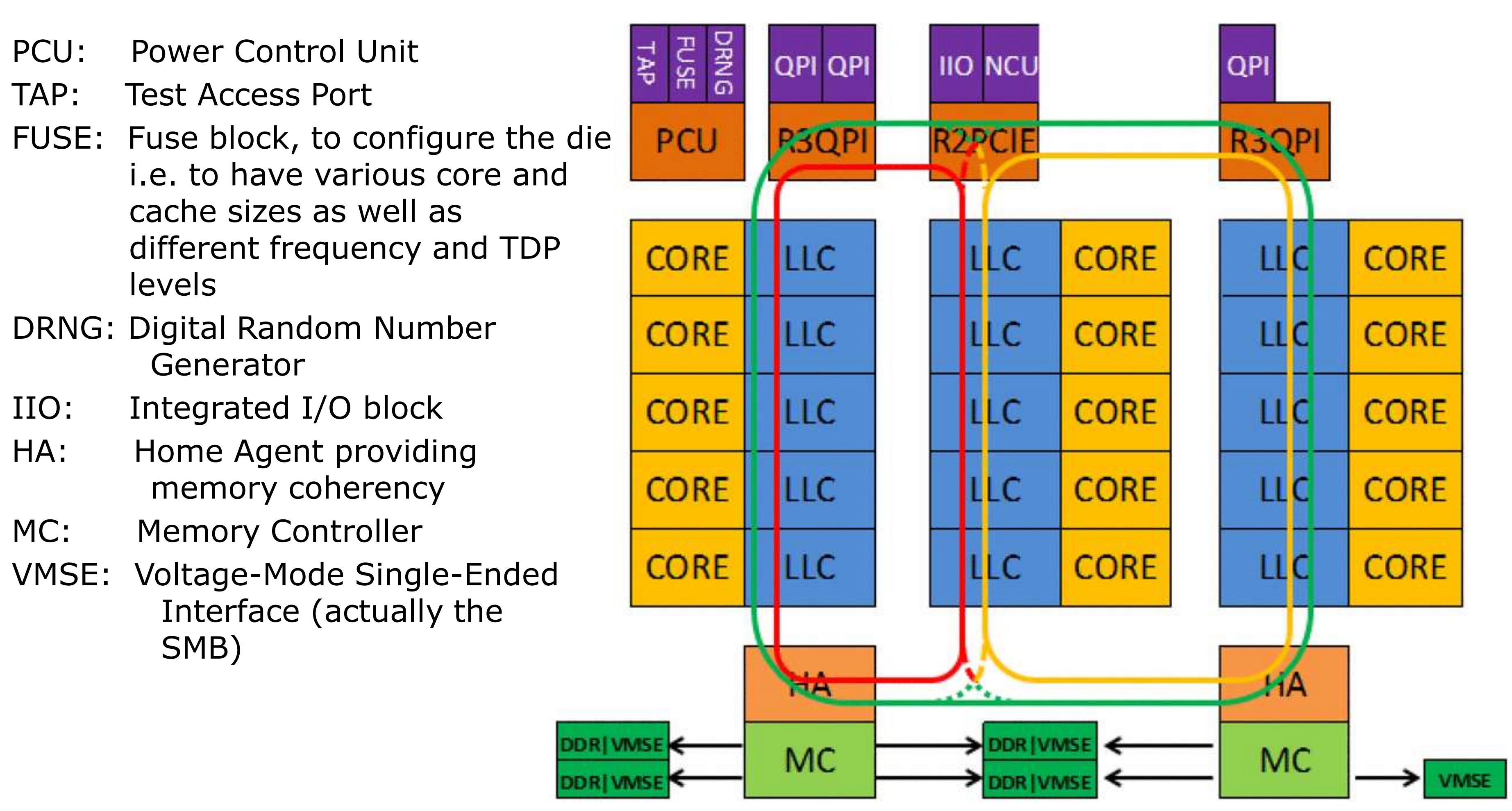
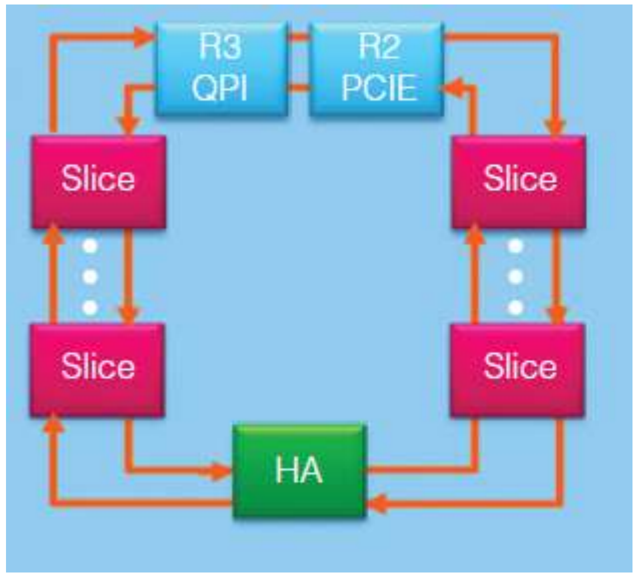


Figure: Block diagram of the Ivy Town processor [116]

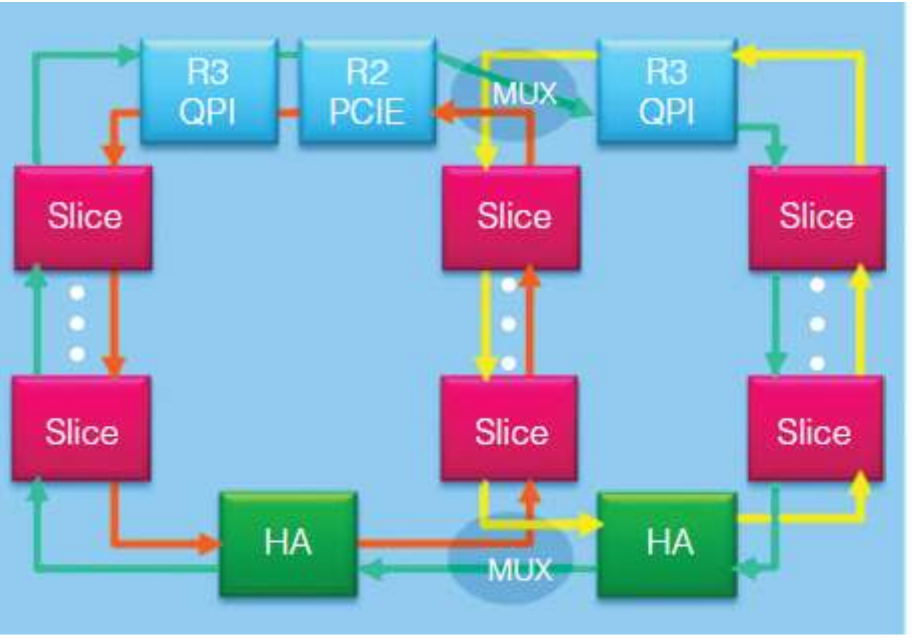
Layout of the ring interconnect bus used for dual slices (10 cores) [125]



Layout of the ring interconnect bus used for three slices (15 cores) [125]

- There are 3 virtual rings.
- Multiplexers (MUXs) dynamically interconnect the rings, as shown below.

Clockwise outer ring



Counter-clockwise outer ring

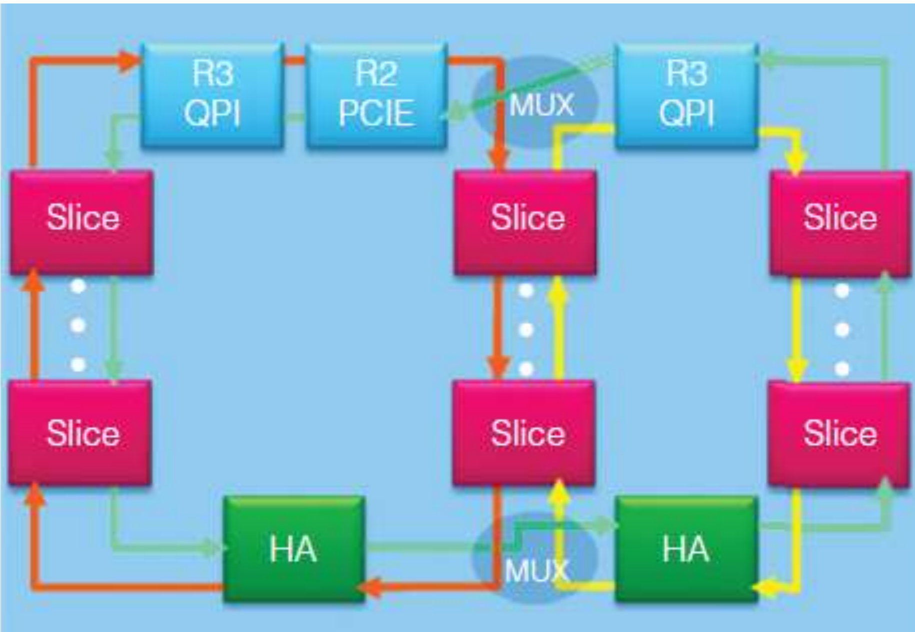
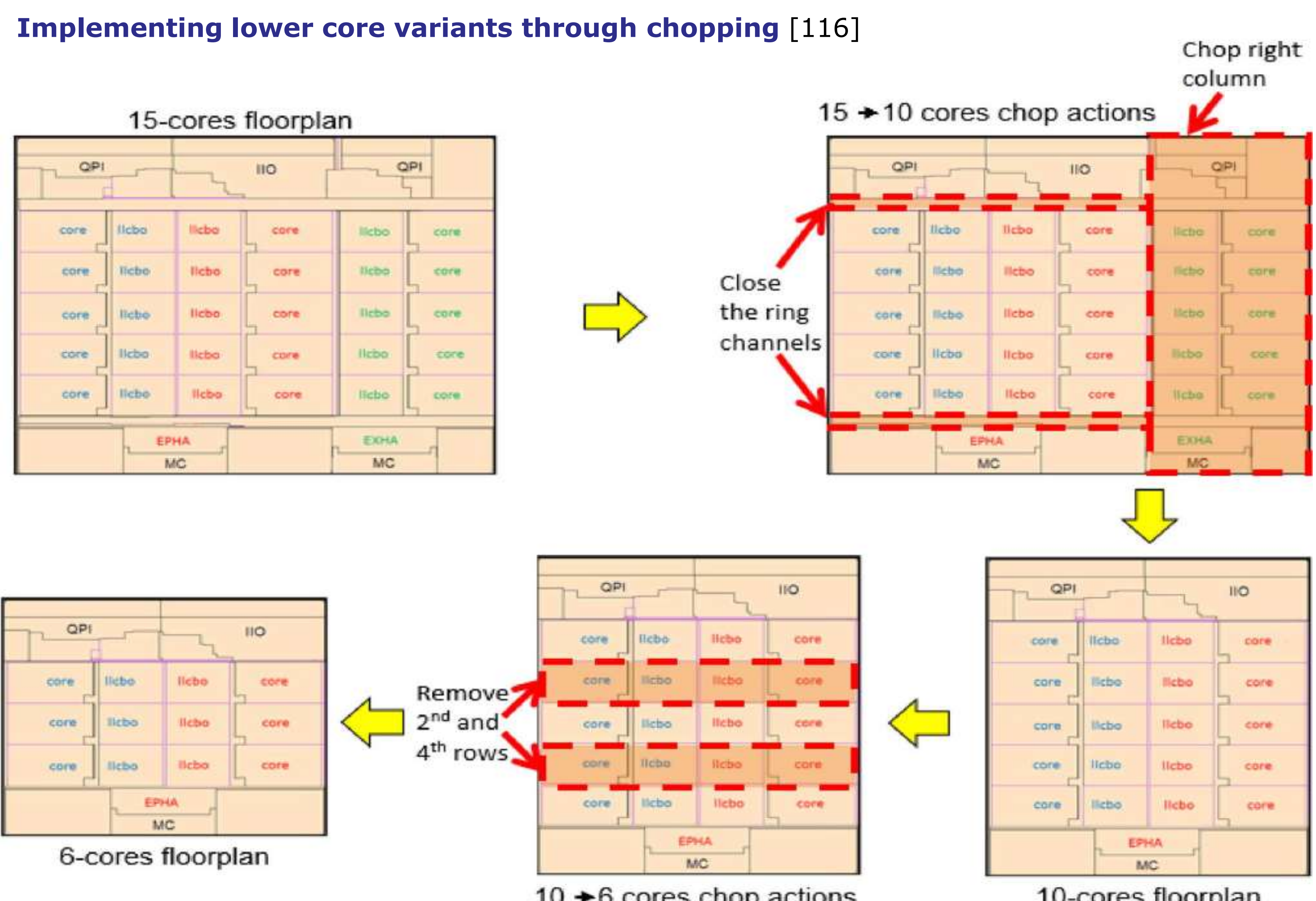
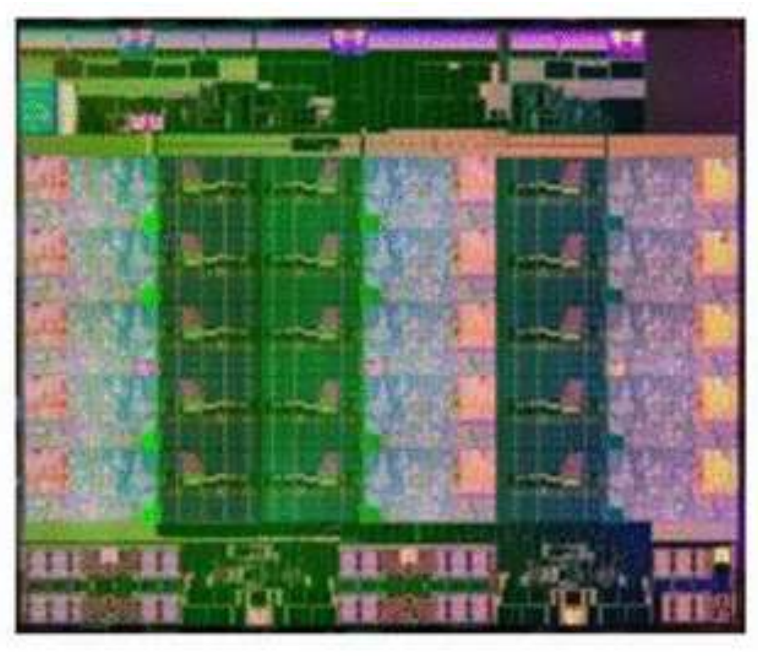
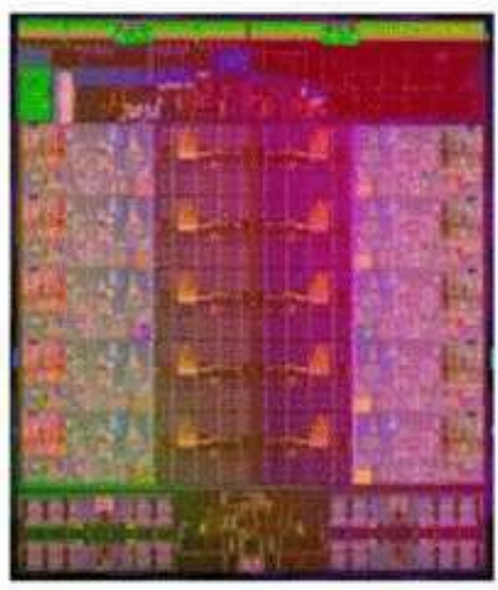
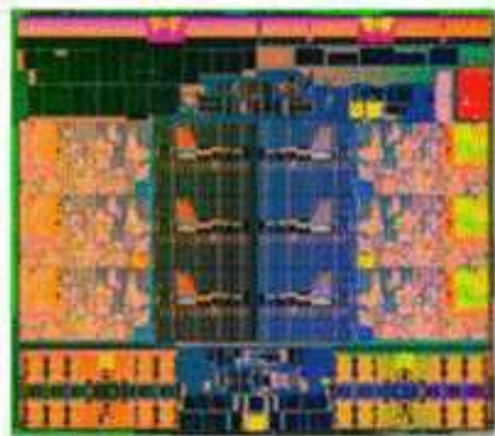


Figure: Three-slice ring interconnect with three virtual rings interconnected by multiplexers [125]

Implementing lower core variants through chopping [116]

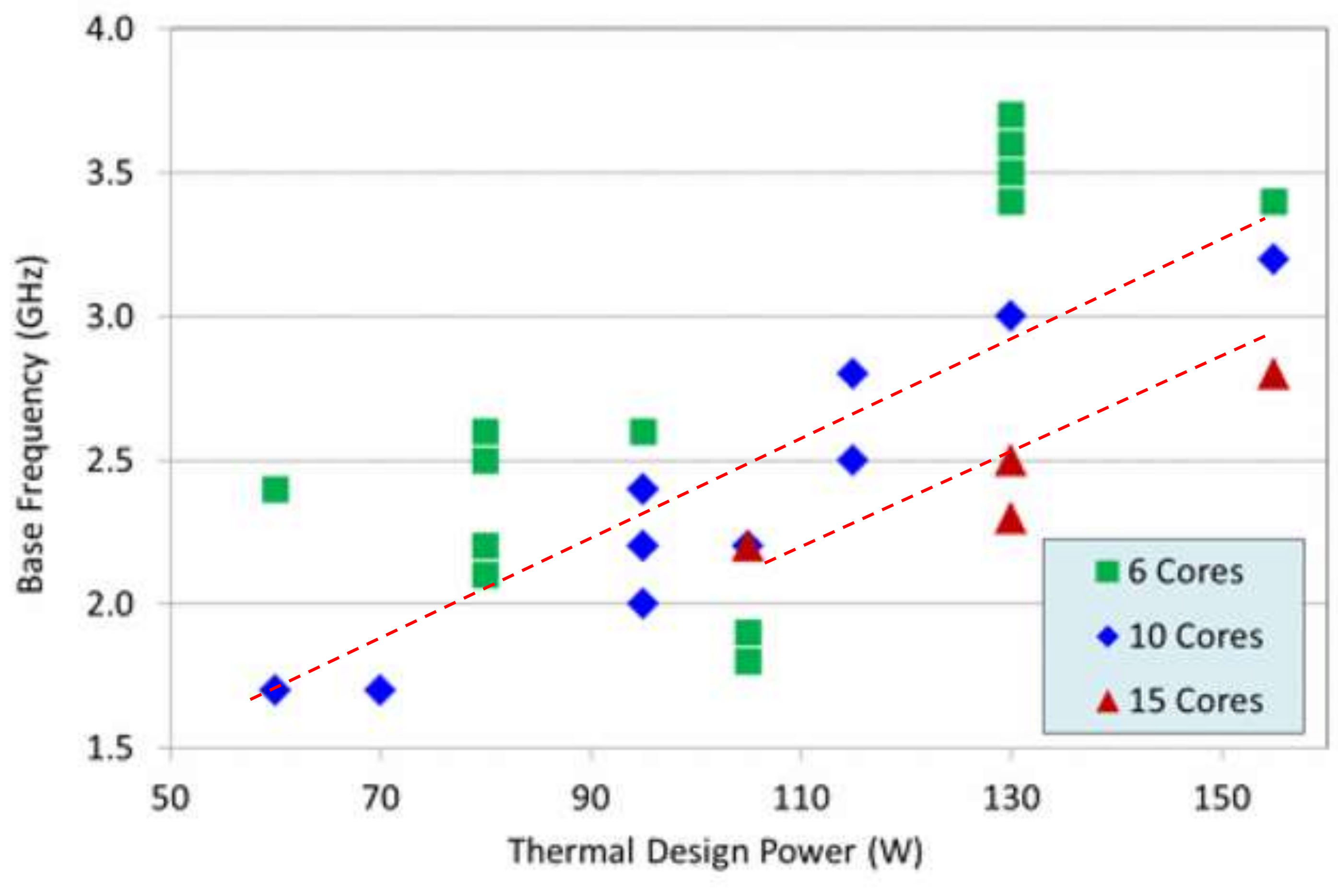


Lower core alternatives of the Ivytown processor achieved by chopping [116]

15 cores	10 cores	6 cores
		

Cores:	15	10	6
L3 cache [MB]:	37.5	25	15
Transistors [B]:	4.31	2.89	1.86
Die size [mm^2]:	541	341	257

TDP vs. core frequency in different Ivytown processor alternatives [116]



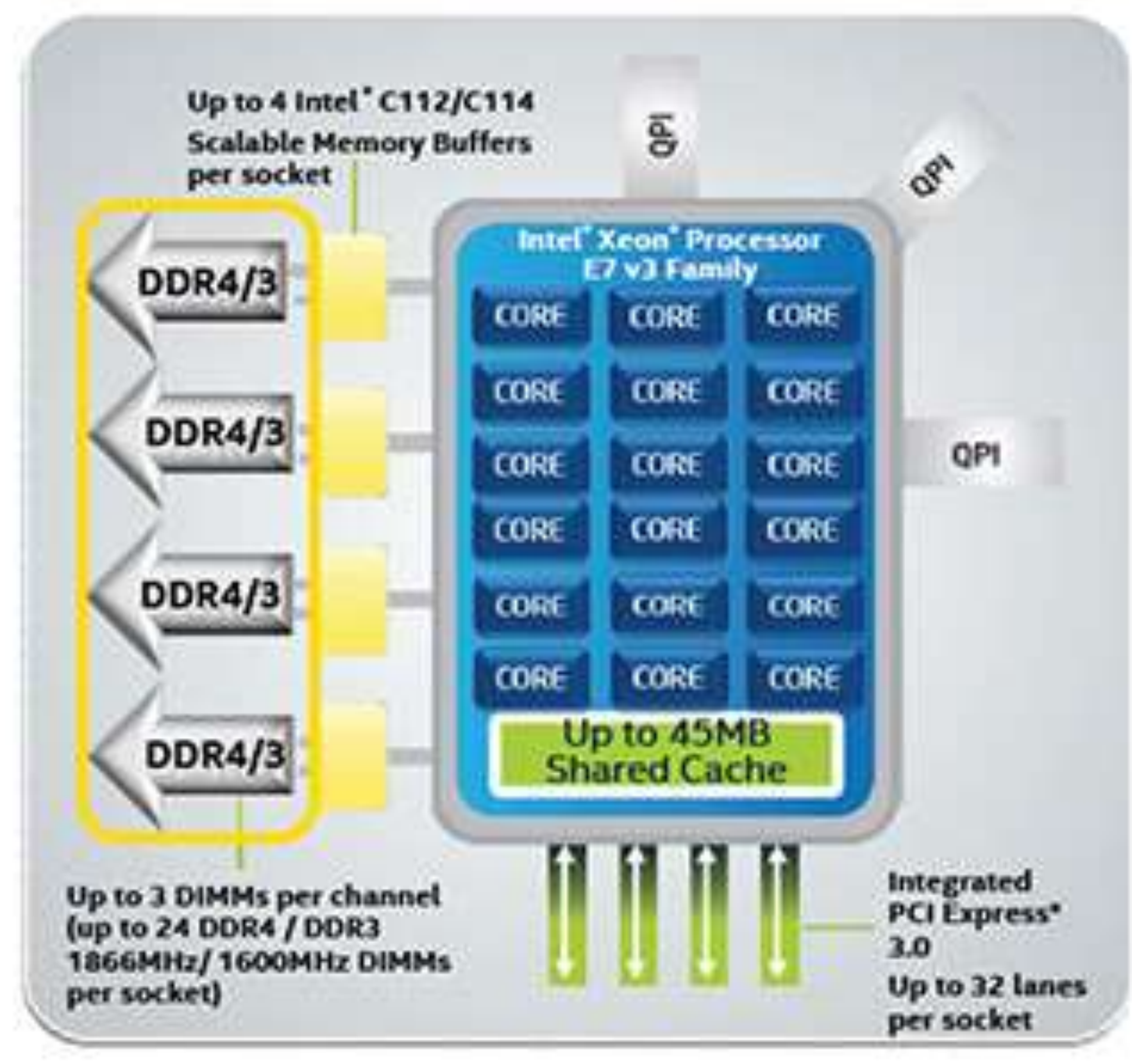
Fuse options enable multiple core and cache sizes (a), as well as different frequency and power levels (b).

3.4 The Haswell-EX (E7-8800 v3) 4S/8S processor line

3.4 The Haswell-EX (E7-8800 v3) 4S/8S processor line

- Launched in 5/2015
- 22 nm, 5.7 billion transistors, 662 mm (for 14 to 18 cores)
- Number of cores
  - 18 cores for 8S processors
  - 14 cores for 4S ones (in 08/2015) instead of 18 cores

Basic layout of an 18 core 8S Haswell-EX v3 (E7-8800) processor [118]



Up to **20%** more cores/threads

Up to **20%** more last level cache for up to **39%** OLTP database performance increase gen-gen[1]

Contrasting key features of the Ivy Bridge-EX and Haswell-EX processors -1 [114]

Feature	Ivy Bridge-EX (E7-8800 v2)	Haswell-EX (E7-8800 v3)
Cores	Up to 15	Up to 18
LLC size (L3)	Up to 15x2.5 MB	Up to 18x2.5 MB
QPI	3xQPI 1.1 Up to 8.0 GT/s in both dir. (16 GB/s per dir.)	3xQPI 1.1 Up to 9.6 GT/s in both dir. (19.2 GB/s per dir.)
SMB	C102/C104 (Jordan Creek) C102: 2 DIMMs/SMB C103: 3 DIMMs/SMB	C112/C114 (Jordan Creek 2) C112: 2 DIMMs/SMB C113: 3 DIMMs/SMB
VMSE speed	Up to 2667 MT/s	Up to 3200 MT/s
DDR4 speed	n.a.	Perf. mode: up to 1600 MT/s Lockstep mode: up to 1866 MT/s
DDR3 speed	Perf. mode: up to 1333 MT/s Lockstep mode: up to 1600 MT/s	Perf. mode: up to 1600 MT/s Lockstep mode: up to 1600 MT/s
TSX	n.a.	Supported

Contrasting key features of the Ivy Bridge-EX and Haswell-EX processors -1 [114]

We note that the Haswell-EX processors support Intel's Transactional Synchronization Extension (TSX) that has been introduced with the Haswell core, but became disabled in the Haswell-EP processors due to a bug.

Transactional Synchronization Extension (TSX)

- Intel's TSX is basically an ISA extension and its implementation to allow hardware supported transactional memory.
- Transactional memory is an efficient synchronization mechanism in concurrent programming used to effectively manage race conditions occurring when multiple threads access shared data.

Addressing race conditions

Basically there are two mechanisms to address race conditions in multithreaded programs, as indicated below:

Basic mechanisms to address races in multithreaded programs

Locks

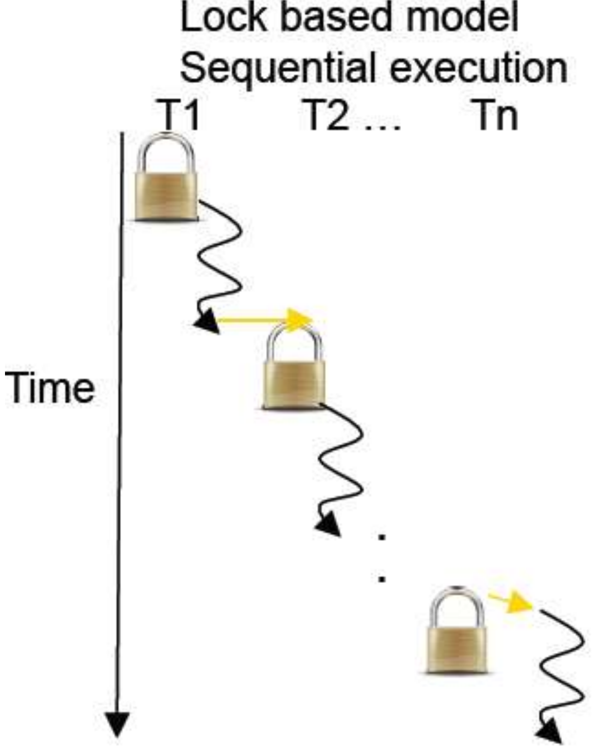
Pessimistic approach,
it intends to prevent possible conflicts
by enforcing serialization of transactions
through locks.

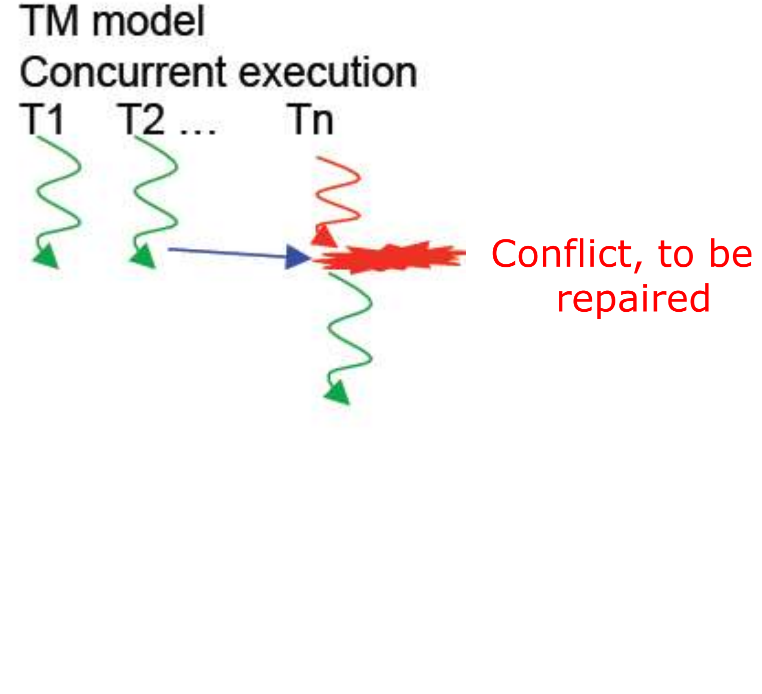
Transactional memory (TM)

Optimistic approach,
it allows access conflicts to occur
but provides a checking and repair mechanism
for managing these conflicts, i.e.
it allows all threads to access shared data simultaneously
but after completing a transaction,
it will be checked whether a conflict arose,
if yes, the transaction will be rolled back and
then replayed if feasible else
executed while using locks.

The next Figure illustrates these synchronization mechanisms.

Illustration of lock based and transaction memory (TM) based thread synchronization [126]





3.4 The Haswell-EX (E7-8800 v3) 4S/8S processor line (8)

We note that in their POWER8 (2014) IBM also introduced hardware supported transactional memory, called Hardware Transactional Memory (HTM).

Contrasting the layout of the Haswell-EX cores vs. the Ivy Bridge-EX cores [111]

Haswell-EX has four core slides with 3x4 + 6 = 18 cores rather than 3x5 cores in case of the Ivy Bridge-EX (only 3x 4 = 12 cores indicated in the Figure below).

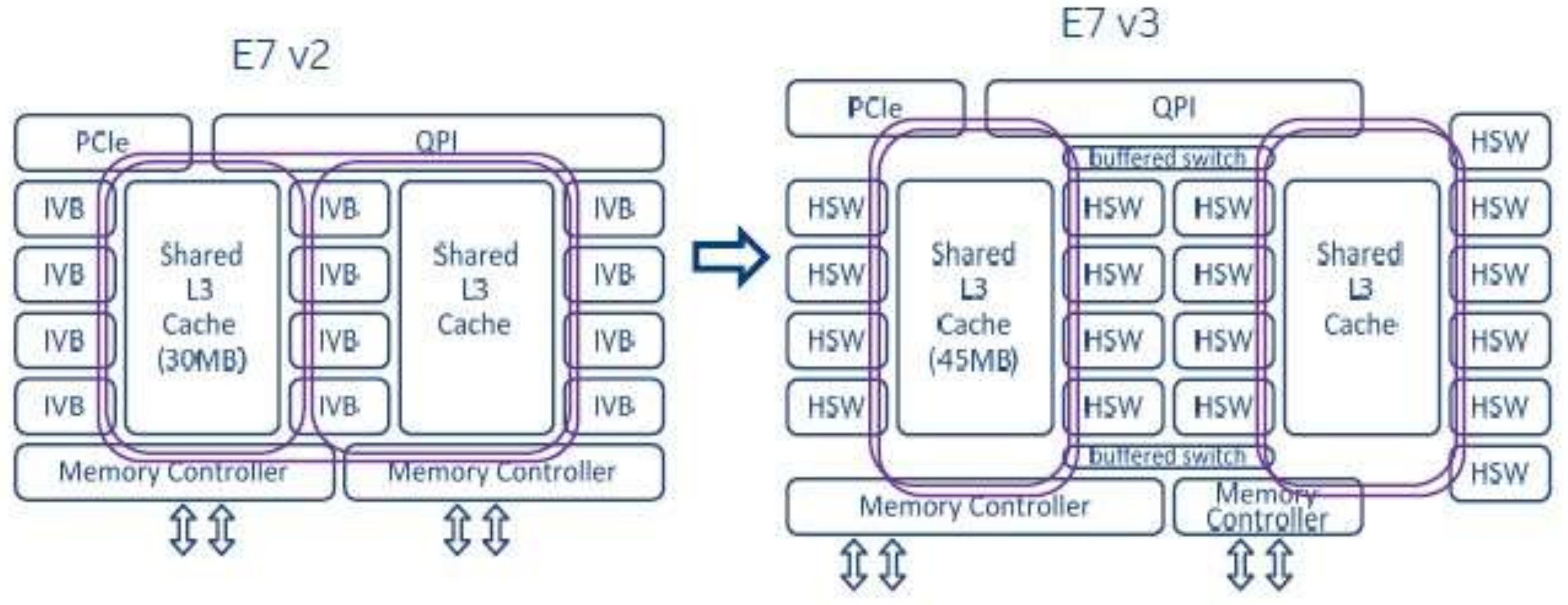
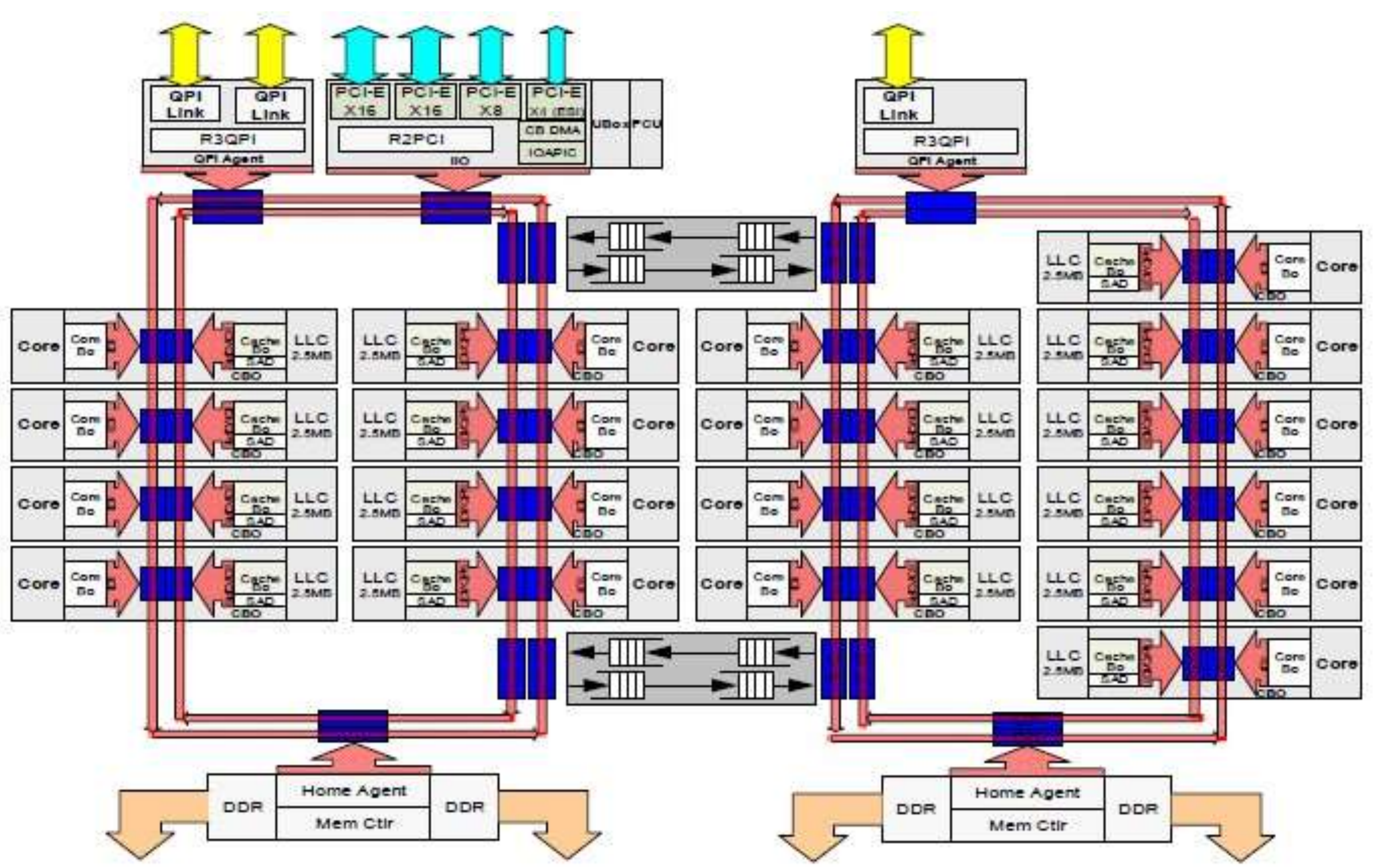


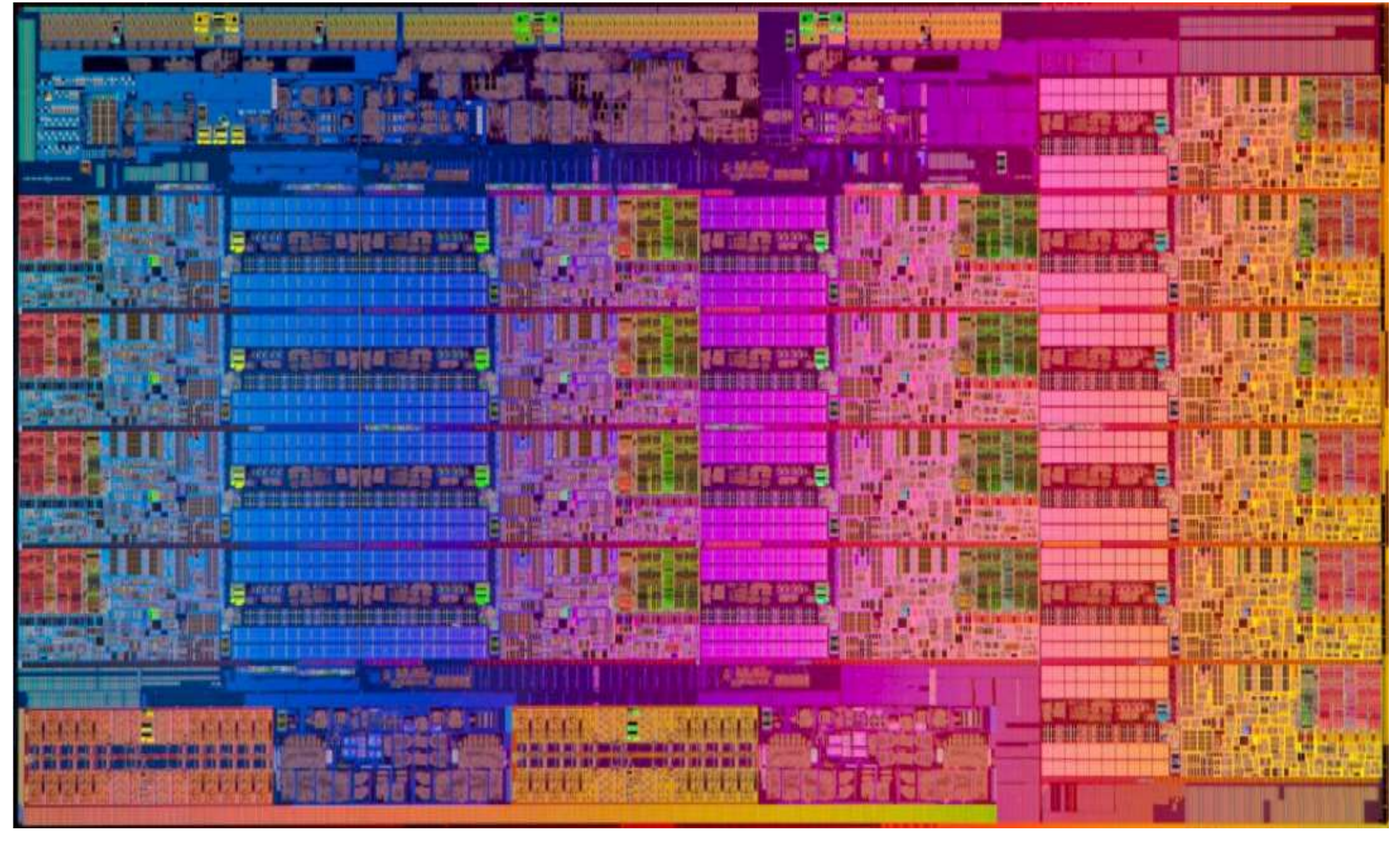
Figure: Contrasting the basic layout of the Haswell-EX (E7 v3) and Ivy Bridge-EX (E7 v2) processors [111]

Note that the E7-V3 has only two ring buses interconnected by a pair of buffered switches.

More detailed layout of the Haswell-EX die [111]



Die micrograph of an 18-core Haswell-EX processor [111]



22 nm, 5.7 billion transistors, 662 mm^2

Main features of 4S and 8S Haswell-EX processors [120]

INTEL® XEON® PROCESSOR E7 v3 FAMILY SPECIFICATIONS

Intel® Xeon® processor SKU	Frequency (GHz)	Cache	Power	Cores	Intel® Turbo Boost Technology	Intel® HT Technology	Intel® QPI Link Speed
E7-8893 v3	3.2	45M	140 W	4	●	●	9.6 GT/s
E7-8891 v3	2.8	45M	165 W	10	●	●	9.6 GT/s
E7-8880L v3	2.0	45M	115 W	18	●	●	9.6 GT/s
E7-8867 v3	2.5	45M	165 W	16	●	●	9.6 GT/s
E7-8890 v3	2.5	45M	165 W	18	●	●	9.6 GT/s
E7-8880 v3	2.3	45M	150 W	18	●	●	9.6 GT/s
E7-8870 v3	2.1	45M	140 W	18	●	●	9.6 GT/s
E7-8860 v3	2.2	40M	140 W	16	●	●	9.6 GT/s
E7-4850 v3	2.2	35M	115 W	14	●	●	8.0 GT/s
E7-4830 v3	2.1	30M	115 W	12	●	●	8.0 GT/s
E7-4820 v3	1.9	25M	115 W	10	No Turbo	●	6.4 GT/s
E7-4809 v3	2.0	20M	115 W	8	No Turbo	●	6.4 GT/s

Power management improvements in the Haswell-EP and Haswell-EX server lines

- The related product families are: the Xeon E5-4600/2600/1600 and the Xeon E7-8800/4800
- The introduced key power management improvements include first of all:
  - Per core P-states (PCPS) and
  - Uncore frequency scaling
- We underline that the installed OS has to support the above features.

Per core P-states (PCPS) in the HSW-EP and Haswell-EX lines [129], [130]

- In the Haswell-EP and -EX server lines Intel introduced individual P-state control instructed by the OS.

  This means that each core can operate independently at individual clock rate and ssociated core voltage.
- Previously, in the Ivy Bridge and preceding generations all cores run at the same frequency and voltage.

3.4 The Haswell-EX (E7-8800 v3) 4S/8S processor line (15)

Remark 1 - Estimated gain of using individual P-states -1

- Subsequently, we cite a study investigating the possible gain of using individual voltage and clock domains [131].
- The assumed microarchitecture of the study is seen below.

PMU: Power Management Unit
Based on sensed data it determines individual clock and voltage values.

Note that due to the different clock rates FIFO Buffers are needed between the cores and the Interconnect for clock synchronization.

Clock synchronization adds latency to the memory and cache transfers

Figure: Assumed microarchitecture in the cited study (clock generation not indicated)

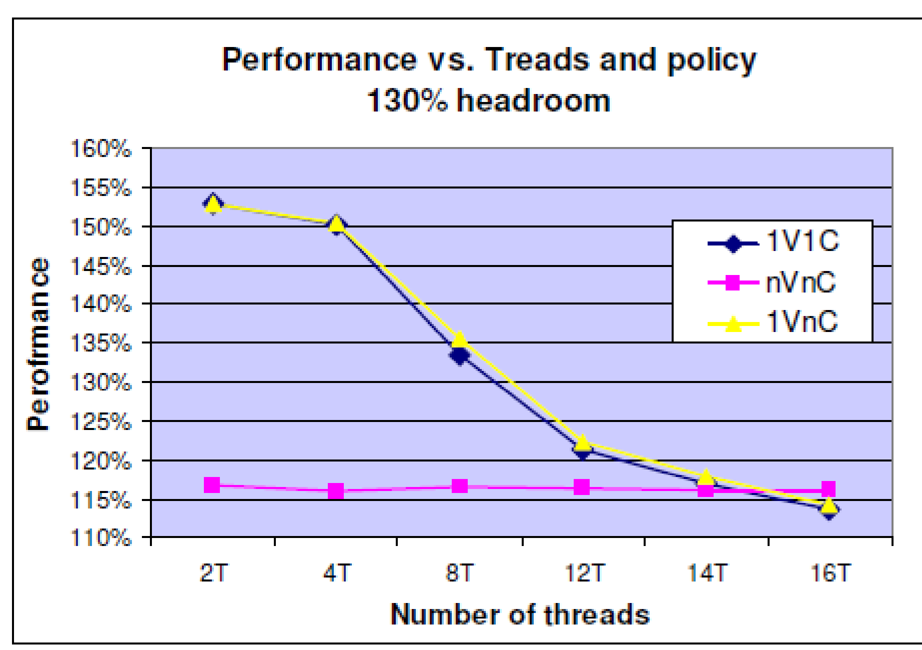
Remark 1 - Estimated gain of using individual P-states -2

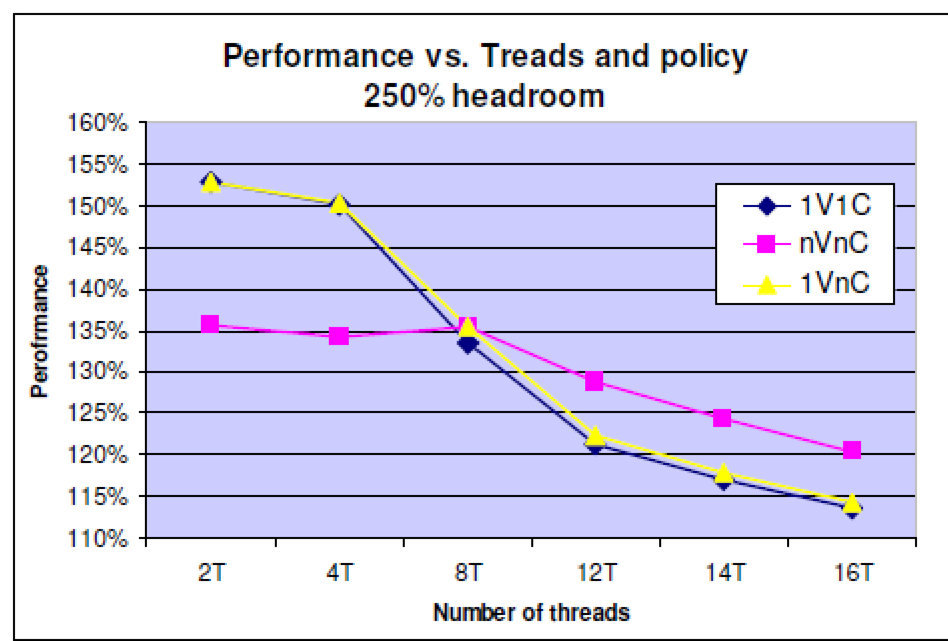
General assessment of using single and multiple voltage domains stated in the study [131]:

"The advantage of a single voltage domain is the capability of sharing the current among the cores; when some cores consume less current or are turned off, current can be directed to the other active cores. This advantage comes at the cost of tying all the voltage domains together, forcing the same operation voltage to all cores.

On the other hand, multiple voltage domains topology provides the ability to deliver individual voltages and frequencies according to an optimization algorithm. In particular, when a single thread workload is executed, the entire CMP power budget can be assigned to a single core, which can consume 16 times higher power than each individual cores when executing a balanced workload on all 16 cores. While in both cases the total CMP power is the same, separate power domains require at least one of the 16 VRs to deliver 16 times higher power than its nominal working point. Such a requirement is not feasible. The present study evaluates VR designed to deliver 130%–250% of the rated CMP current."

Estimated gain vs a baseline platform that does not make use of DVFS [131]





1V1C: Single Voltage domain, single Clock domain
nVnC: Multiple Voltage domains, multiple Clock domains
1VnC: Single Voltage domain, multiple Clock domains

The headroom is understood as the capability of the voltage regulators to deliver up to 130 % or 150 % of the rated current (the current consumed at the min. Vcc design point (the lowest P point).

- The study has shown that for low thread counts, typical in DT and mobile platforms, the use of multiple voltage and clock domains degrades performance due to clock synchronization needed that adds latency to memory and cache transfers.
- By contrast, in processors that support a large number of threads, as typical in servers, multiple voltage and frequency domains may be benefitial assuming that voltage regulators have a high enough power delivery capability.

Remark 2 -Principle of operation of the Per-core P-state (PCPS) power management (simplified) -1

- US patent application 20140229750 A1 (filed in 2012 and issued in 2014) reveals details of the implementation [133].
- The Figure below illustrates PCPS for an 18 core Haswell-EX processor when the cores run a four different P-states.

Method an
US 20140

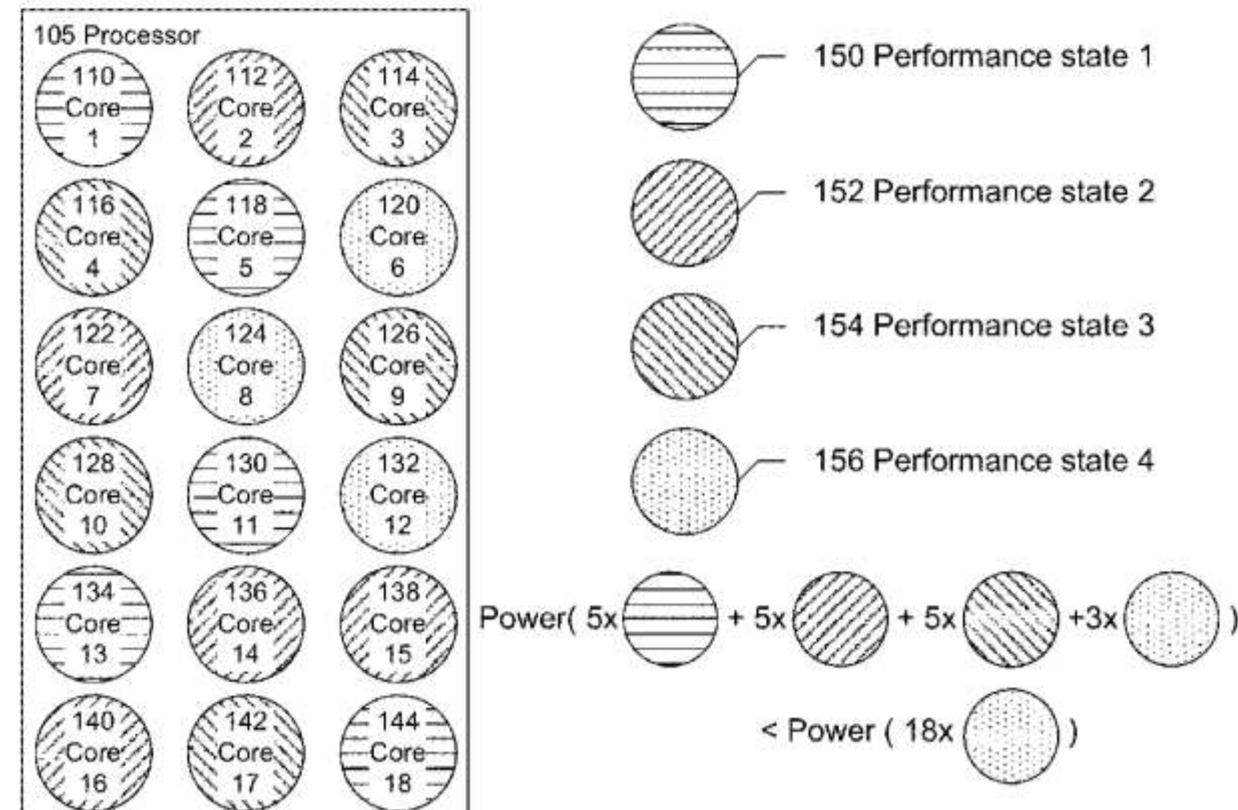


Figure: Illustration of PCPS for an 18-core Haswell-EX processor if the cores run a four P-states [133

Remark 2 -Principle of operation of the Per-core P-state (PCPS) power management (simplified) -2

- According to the patent each core includes a set of 32-bit registers, one register for each thread (Registers 212 to 214) plus an additional register (Register 218), as indicated for two cores (Core i and Core n) below.

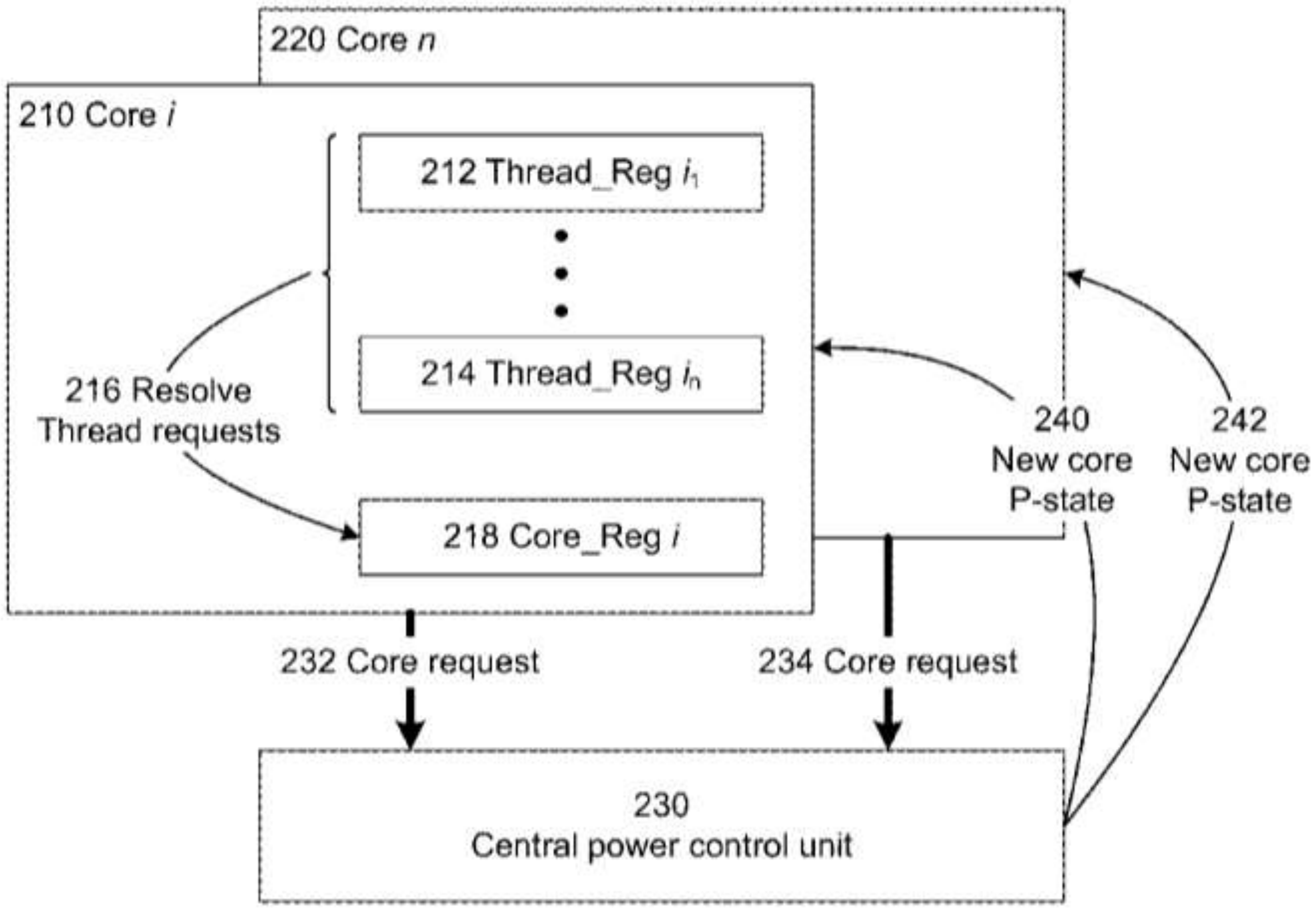


Figure: Per core implemented registers used for PCPS [133]

Remark 2 -Principle of operation of the Per-core P-state (PCPS) power management (simplified) -3

- The per-thread available registers (Registers 212 to 214) provide a set of fields, detailed in the patent.
- The OS determines for each thread the requested P-state needed for executing the thread at efficient performance (i.e. at the min. clock rate needed to execute the thread without substantially lengthening its runtime. Note here that each thread is scheduled only for in given time windows thus a higher clock rate would not substantially reduce its runtime).
- Available core logic evaluates the per-thread requests, determines the highest performance P-state needed for the core and writes this values into the additional register (Register 218).
- Finally, the requested P-state is communicated to the Central Power Control Unit (230) and it sets the related individual voltage regulator and clock generator (PLL) belonging to the core considered to obtain the requested P-state.

Uncore frequency scaling [132]

- With uncore frequency scaling the uncore voltage and frequency is independently controlled from the core's voltage and frequency settings.

  To achieve this the uncore needs to be implemented on a separate voltage and clock domain.

- Benefit of uncore frequency scaling
  - For compute bound applications core frequency may be raised without needing to increase uncore frequency and voltage or
  - for memory-bound applications uncore frequency may be raised without needing to increase core frequency and voltage.

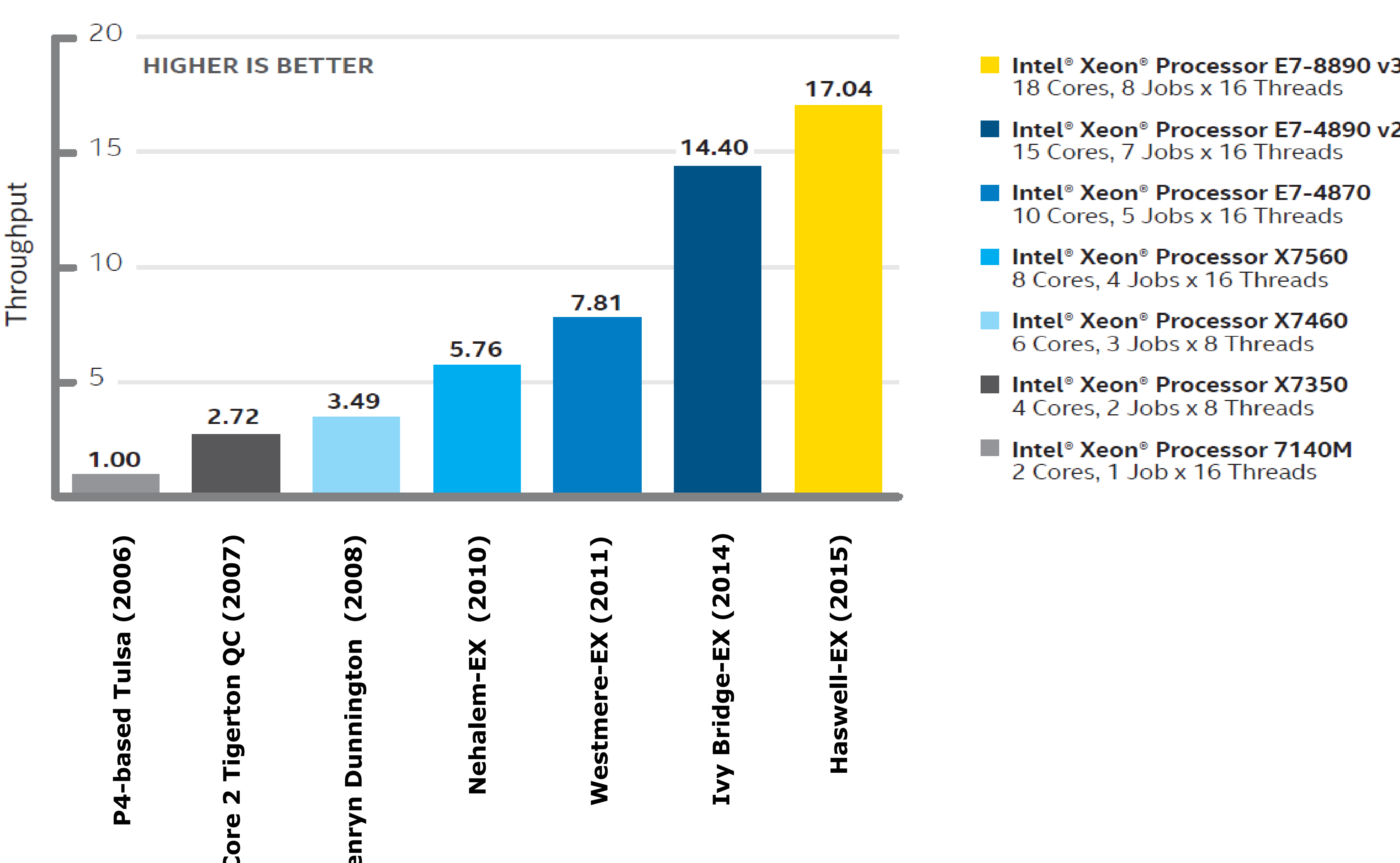
  This allows to optimize performance by consuming less power.

Remark [132]

- In Nehalem the cores were DVFS controlled whereas the uncore run at a fixed frequency
- In the Sandy Bridge and Ivy Bridge cores and uncore were tight together
- in Haswell each core and the uncore is treated separately.

Performance comparison of high-end 4S/8S server processors -1 [121]

(The test workload is: CPU 45 nm Design Rule Check Tool C 2006 Rev.)



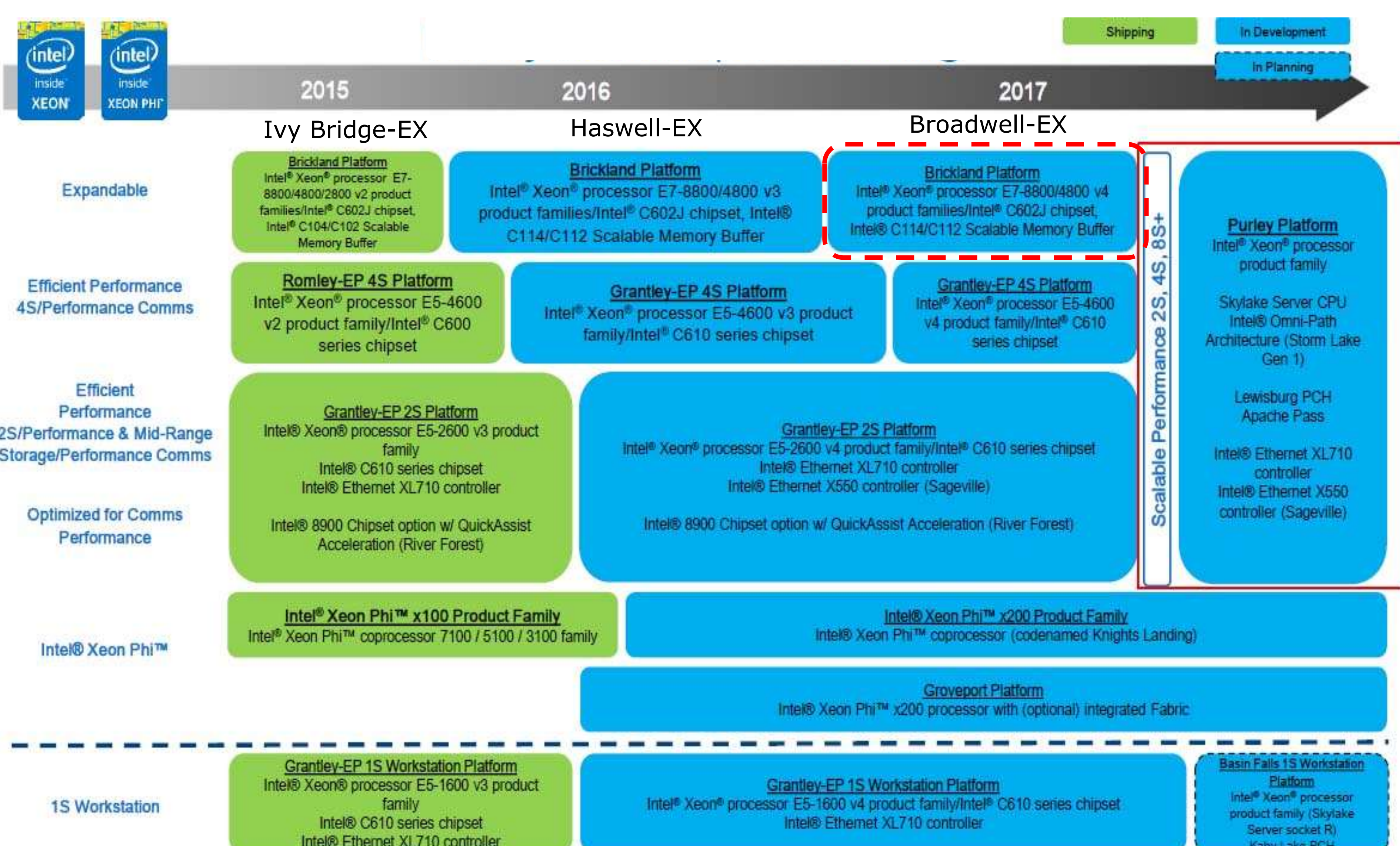
Performance comparison of Intel's high-end 4S/8S server processors -2 [121]

Note that with their high-end 4S/8S server processors Intel achieved a more than 10-fold performance boost in 10 years.

3.5 The Broadwell-EX (E7-8800 v4) 4S/8S processor line

3.5 The Broadwell-EX (E7-8800 v4) 4S/8S processor line

Positioning the 14 nm Broadwell-EX line [112]



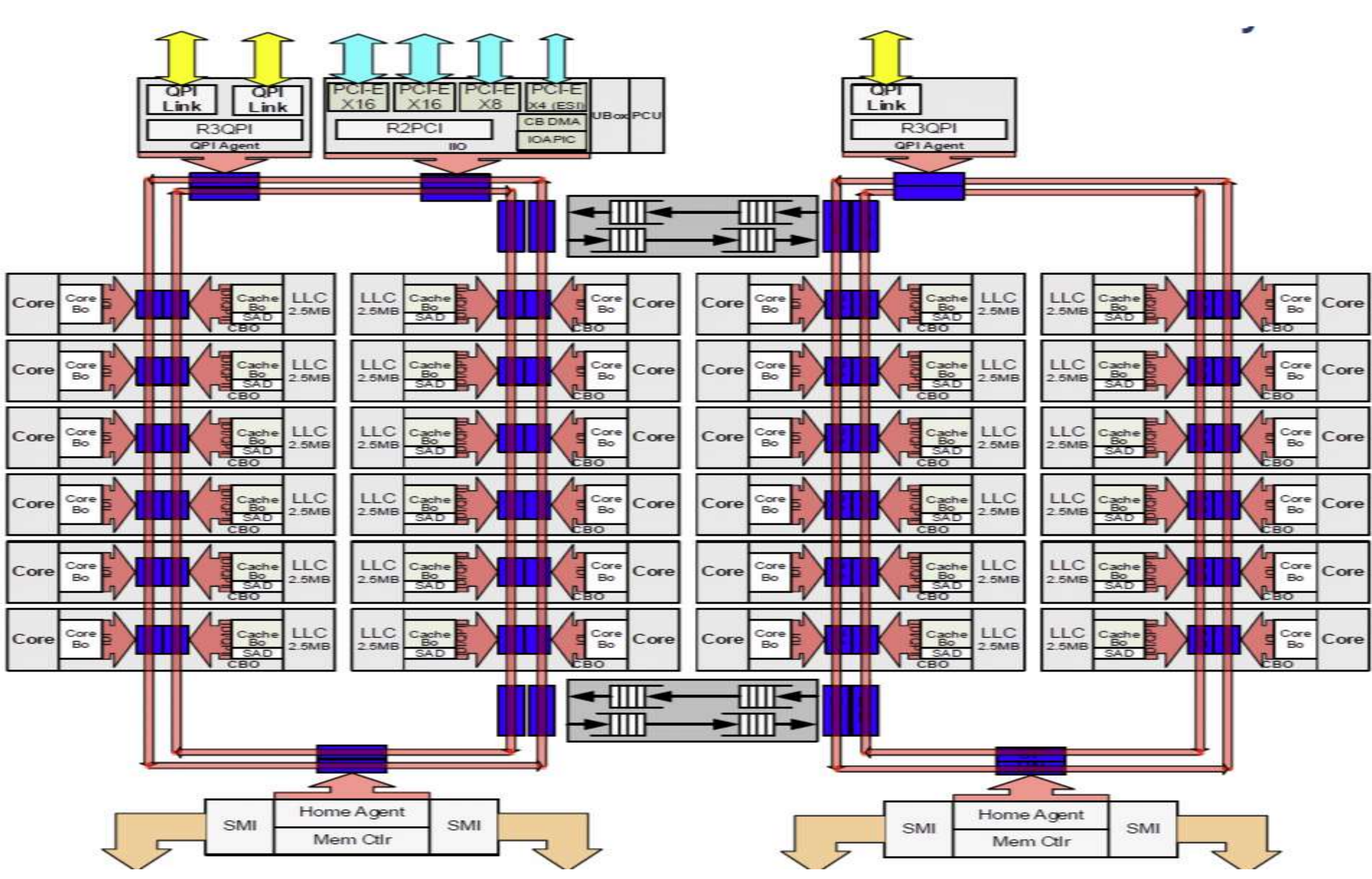
Planned features of the Brickland platform and the Broadwell-EX processor line -1 [122]

Spec	Grantley with Broadwell-EP CPU	Brickland with Broadwell-EX CPU	Purley with Skylake CPU
CPU TDP (with IVR)	55-145W, 160W WS only	115-165W	45-165W
Socket	Socket R3	Socket R1	Socket P
Scalability	2S	2S, 4S, 8S	2S, 4S, 8S
Cores	Up to 22C with Intel® HT Technology	Up to 24C with Intel® HT Technology	Up to 28C with Intel® HT Technology
Memory	4 channels DDR4 per CPU RDIMM, LRDIMM	4 channels DDR4 per CPU RDIMM, LRDIMM	6 channels DDR4 per CPU RDIMM, LRDIMM
	1DPC=up to 2400, 2DPC= up to 2133, 3DPC=up to 1600	DDR3/4 Performance Mode 1333, 1600 DDR3/4 Lockstep mode 1333, 1600, 1866	2133, 2400 2DPC, 2666 1DPC No 3 DPC support
UPI	QPI: 2 v1.1 channels per CPU 9.6 GT/s max	QPI: 3 v1.1 channels per CPU 9.6 GT/s max	UPI: 2-3 channels per CPU (9.6, 10.4 GT/s)
PCIe*	PCIe* 3.0 (2.5, 5.0, 8.0 GT/s)	PCIe* 3.0 (2.5, 5.0, 8.0 GT/s)	PCIe* 3.0 (2.5, 5.0, 8.0 GT/s)
	40 lanes per CPU	32 lanes per CPU	48 lanes per CPU Bifurcation support: x16, x8, x4
PCH	Wellsburg: DMI2 – 4 lanes; Up to 6xUSB3, 8x USB2 ports, 10xSATA3 ports; GbE MAC (+ External PHY)	Patsburg: 14 USB2 ports, 4 SATA2 ports, 2 SATA3 ports	Lewisburg: DMI3 – 4 lanes; 14xUSB2 ports Up to: 10xUSB3; 14xSATA3, 20xPCIe*3 New: Innovation Engine, 4x10GbE ports, Intel® QuickAssist Technology
External Node Controller Support	None	3rd Party Node Controller	3rd Party Node Controller supported on select skus

Main features of the Broadwell-EX processor line -2 [122]

As seen in the above Table, the planned features – except of having 24 cores – are the same as implemented in the Haswell-EX line.

Layout of the Broadwell-EX die [128]



4S Cluster on Die (COD) Mode [128]

- COD is a performance enhancing feature, introduced with Haswell-EP, as 2S COD mode.
- Broadwell-EX adds 4S COD capability by dividing 24 cores, 24 LLCs and the Home Agents (HA) into two clusters, as seen below.

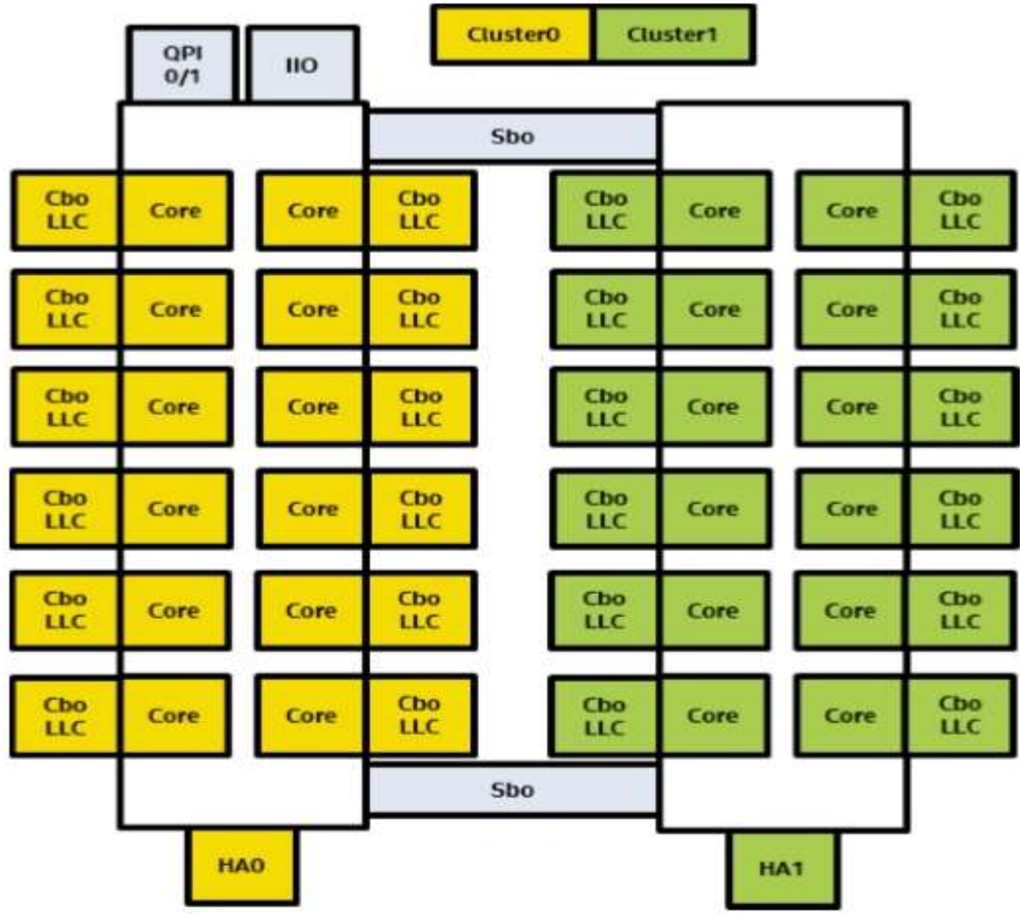


Figure: 4S Cluster on Die (COD) Mode [128]

3.5 The Broadwell-EX (E7-8800 v4) 4S/8S processor line (6)

Principle of operation [128]

- Each LLC cluster operates as an independent caching agent.
- OS creates NUMA domains such that most memory accesses remain within the cluster.

Benefits

- LLC access latency is reduced as the cache slices are more localized to the cores.
- Memory system latency is reduced because the number of threads seen by each Home Agent is reduced and this increases the likelihood that memory requests hit open pages in the memory controller.

Remark - Introduction of the COD mode already in Haswell-EP for 1S and 2S servers [134]

In case of an 18 core Haswell-EP (E5-2600 v3) processor the cores are partitioned into 2x 9 cores as follows:

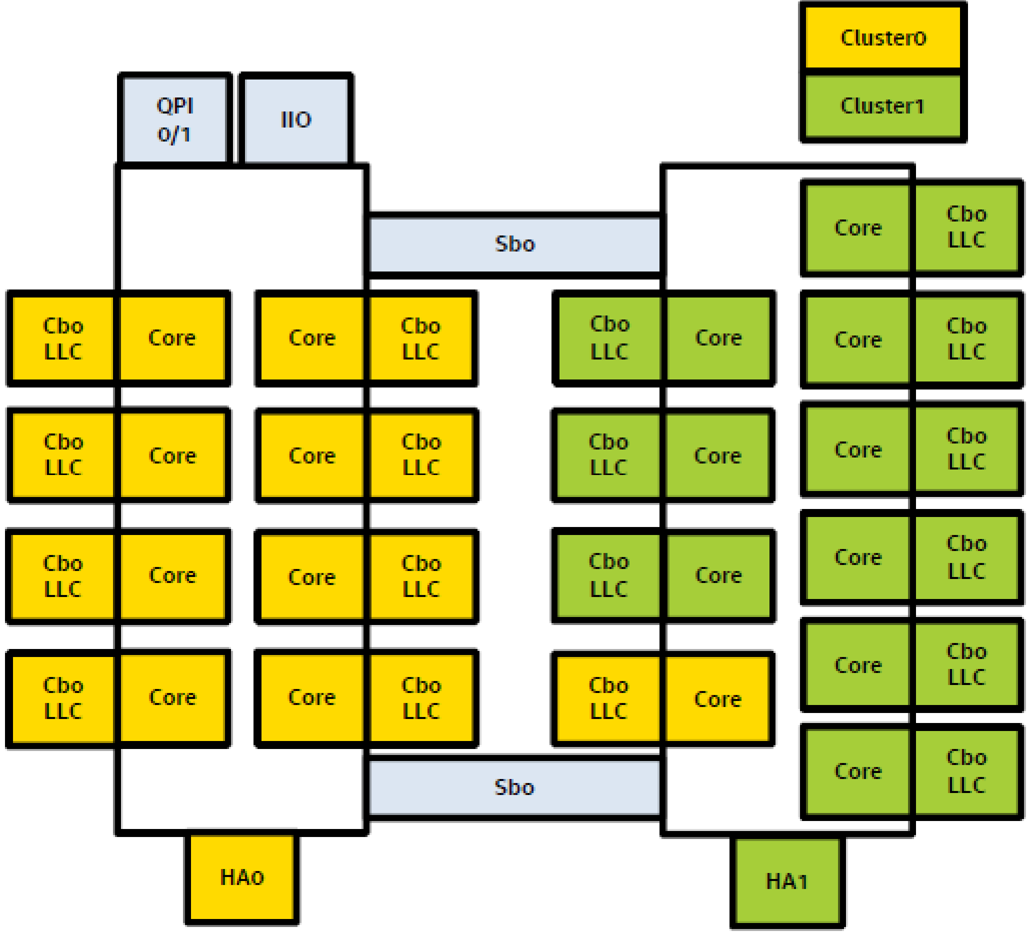


Figure: Partitioning 18 cores into 2x 9 cores for the COD mode in a Haswell-EP processor [134]

Main features of the Broadwell-EX (Xeon E7-8800 v4 and 4800 v4) lines [128]

	Cores / Threads	Clock / Turbo	L3 Cache	QPI	TDP	MSRP
Xeon E7-8890 v4	24 / 48	2.2 / 3.4 Ghz	60 MB	9.6 GT / s	165 W	$7174.00
Xeon E7-8880 v4	22 / 44	2.2 / 3.3 Ghz	55 MB	9.6 GT / s	150 W	$5895.00
Xeon E7-8870 v4	20 / 40	2.1 / 3.0 Ghz	50 MB	9.6 GT / s	140 W	$4672.00
Xeon E7-8860 v4	18 / 36	2.2 / 3.2 Ghz	45 MB	9.6 GT / s	140 W	$4061.00
Xeon E7-8855 v4	14 / 28	2.1 / 2.8 Ghz	35 MB	9.6 GT / s	140 W	N/A
Xeon E7-8867 v4	18 / 36	2.4 / 3.3 Ghz	45 MB	9.6 GT / s	165 W	$4672.00
Xeon E7-8891 v4	10 / 20	2.8 / 3.5 Ghz	60 MB	9.6 GT / s	165 W	$6841.00
Xeon E7-8893 v4	4 / 8	3.2 / 3.5 Ghz	60 MB	9.6 GT / s	140 W	$6841.00
Xeon E7-4850 v4	16 / 36	2.1 / 2.8 Ghz	40 MB	8.0 GT / s	115 W	$3003.00
Xeon E7-4830 v4	14 / 28	2.0 / 2.8 Ghz	35 MB	8.0 GT / s	115 W	$2170.00
Xeon E7-4820 v4	10 / 20	2.0 Ghz	25 MB	6.4 GT / s	115 W	$1502.00
Xeon E7-4809 v4	8 / 16	2.1 Ghz	20 MB	6.4 GT / s	115 W	$1223.00

4. Example 2: The Purley platform

- 4.1 Overview of the Purley platform
- 4.2 Key innovations of the Purley platform vs. the previous Brickland platform
- 4.3 The Skylake-SP (Xeon x100) 2S/4S/8S processor line

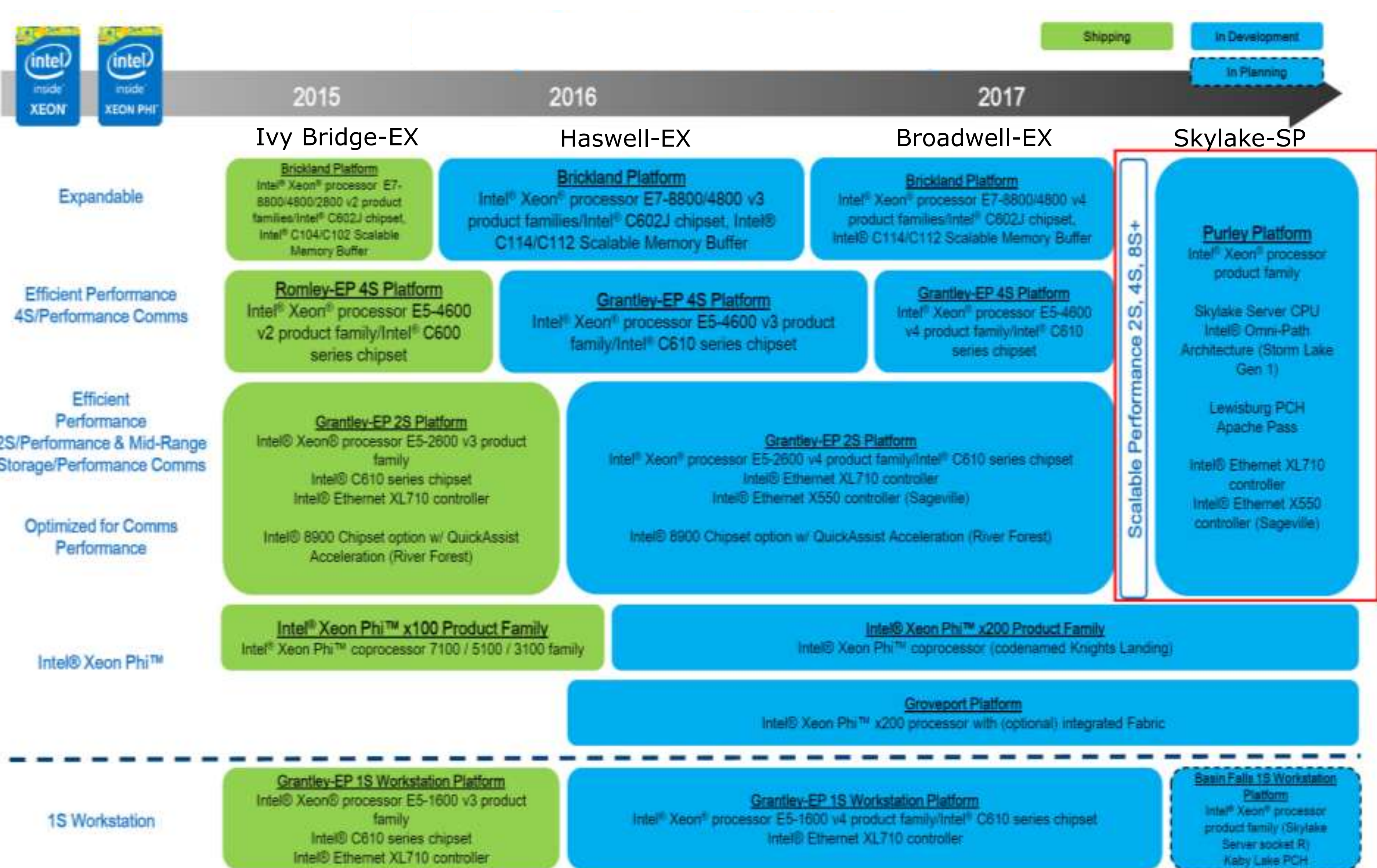
4.1 Overview of the Purley platform

4.1 Overview of the Purley platform

- The Purley platform is introduced in 7/2017, termed also as the Scalable platform.
- It is based on the 14 nm Skylake-SP (Scalable Family Processor) server lines.
- Designed for existing and emerging server applications, such as data centers, cloud computing, HPC and artificial intelligence.

*

Positioning of the Purley platform [112]



4.1 Overview of the Purley platform (3)

Main features of the Purley platform based on the Skylake-SP processor line -1 [122]

Spec	Grantley with Broadwell-EP CPU	Brickland with Broadwell-EX CPU	Purley with Skylake CPU
CPU TDP (with IVR)	55-145W, 160W WS only	115-165W	45-165W
Socket	Socket R3	Socket R1	Socket P
Scalability	2S	2S, 4S, 8S	2S, 4S, 8S
Cores	Up to 22C with Intel® HT Technology	Up to 24C with Intel® HT Technology	Up to 28C with Intel® HT Technology
Memory	4 channels DDR4 per CPU RDIMM, LRDIMM	8 ~~4~~ channels DDR4 per CPU RDIMM, LRDIMM	6 channels DDR4 per CPU RDIMM, LRDIMM
	1DPC=up to 2400, 2DPC= up to 2133, 3DPC=up to 1600	DDR3/4 Performance Mode 1333, 1600 DDR3/4 Lockstep mode 1333, 1600, 1866	2133, 2400 2DPC, 2666 1DPC No 3 DPC support
UPI	QPI: 2 v1.1 channels per CPU 9.6 GT/s max	QPI: 3 v1.1 channels per CPU 9.6 GT/s max	UPI: 2-3 channels per CPU (9.6, 10.4 GT/s)
PCIe*	PCIe* 3.0 (2.5, 5.0, 8.0 GT/s)	PCIe* 3.0 (2.5, 5.0, 8.0 GT/s)	PCIe* 3.0 (2.5, 5.0, 8.0 GT/s)
	40 lanes per CPU	32 lanes per CPU	48 lanes per CPU Bifurcation support: x16, x8, x4
PCH	Wellsburg: DMI2 – 4 lanes; Up to 6xUSB3, 8x USB2 ports, 10xSATA3 ports; GbE MAC (+ External PHY)	Patsburg: 14 USB2 ports, 4 SATA2 ports, 2 (C602J) SATA3 ports	Lewisburg: DMI3 – 4 lanes; 14xUSB2 ports Up to: 10xUSB3; 14xSATA3, 20xPCIe*3 New: Innovation Engine, 4x10GbE ports, Intel® QuickAssist Technology
External Node Controller Support	None	3rd Party Node Controller	3rd Party Node Controller supported on select skus

DPC: DIMMs Per Channel

4.1 Overview of the Purley platform (4)

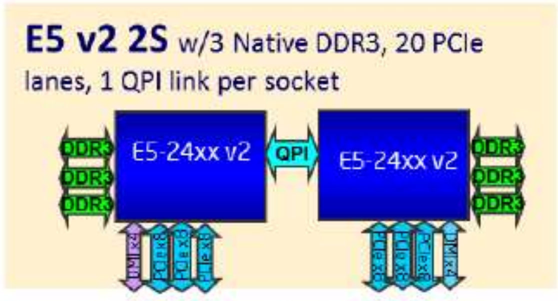
Drawback of Intel's recent server lines

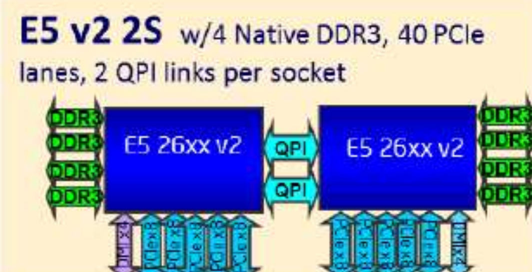
Large diversity of the recent server implementations, as indicated in the next slides.

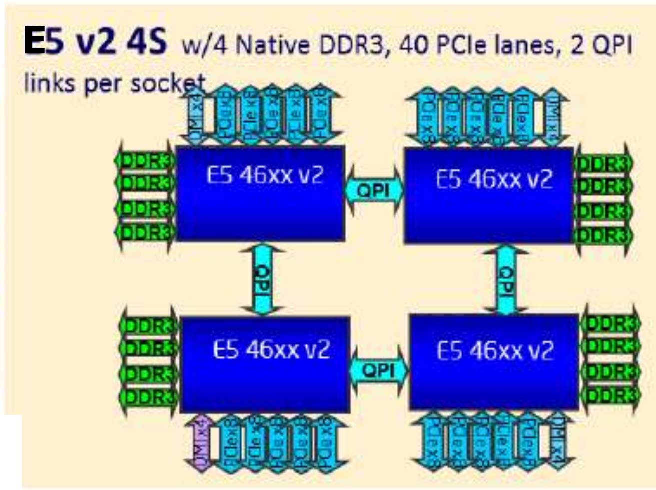
Example: E5/E7 platform options built up of Ivy Bridge based server processors [117]

Romley platform: Entry and Efficient Performance level platform options

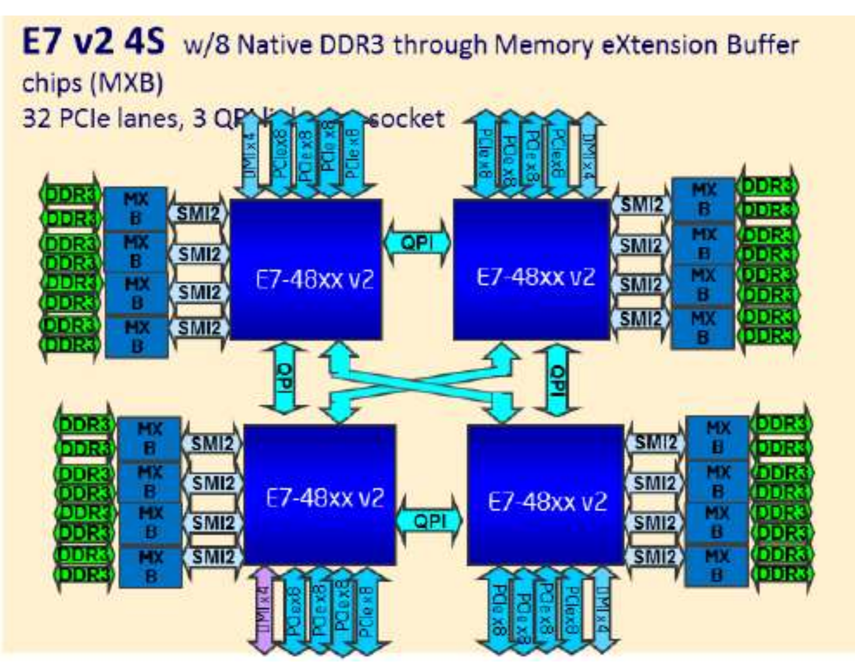
EN: 24xx v2, EP: 26xx v3, 46xx v2

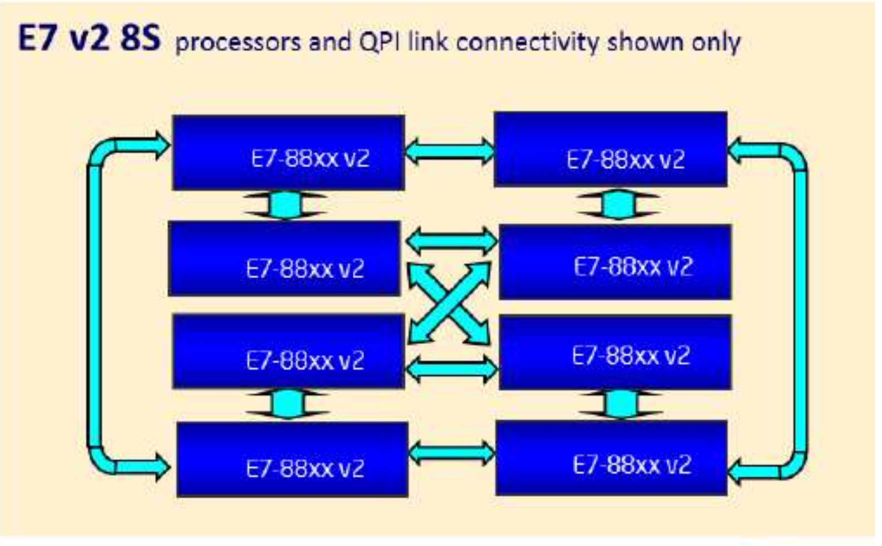






Brickland platform: EX (48xx/88xx v2) platform options





Different kind of attaching memory in EN/EP vs. EX servers (from Nehalem to Broadwell)

Attaching memory to Intel's servers

Direct attaching memory to EN/EP servers

Attaching up to 4 standard DDR memory channels directly to to processor die

Indirect attaching memory to EX servers

Attaching 8 standard DDR memory channels via 4 low line count serial or proprietary 64-bit parallel channels (SMI) with memory buffers, while two standard DDR memory channels can be connected to each memory buffer (SMB)

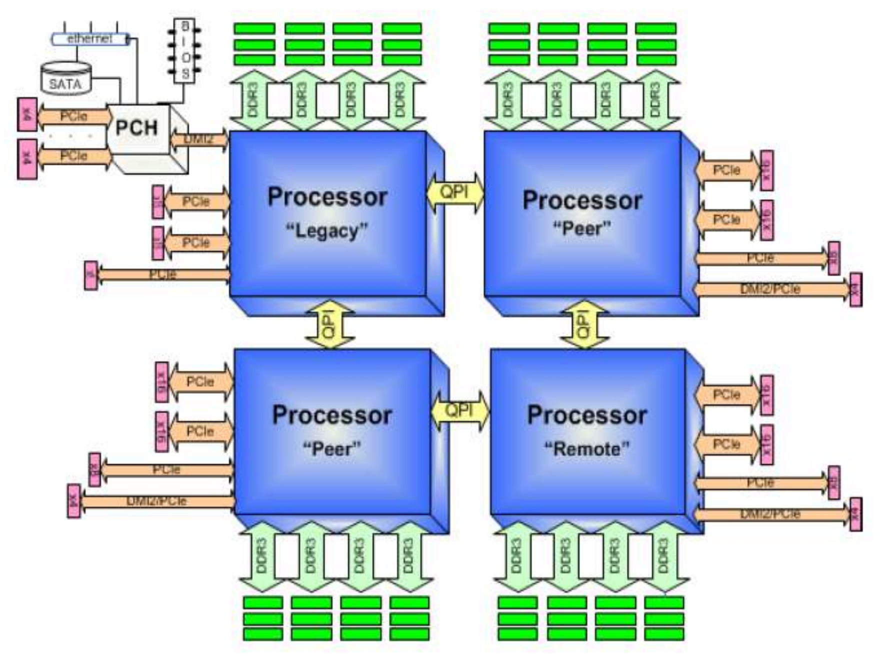


Figure: Sandy Bridge-EP based server [142]

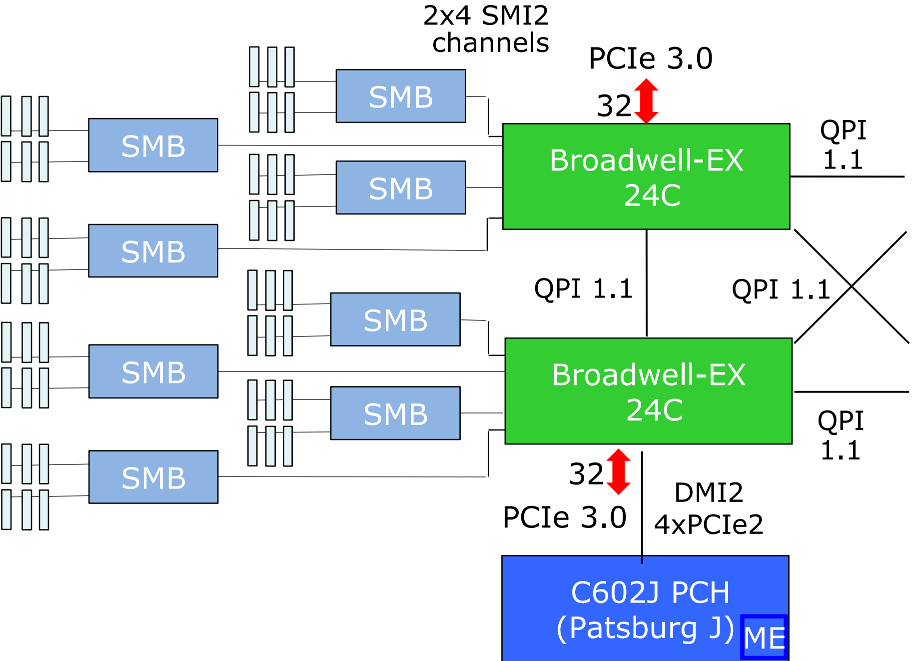


Figure: Part of a Broadwell-EX based server

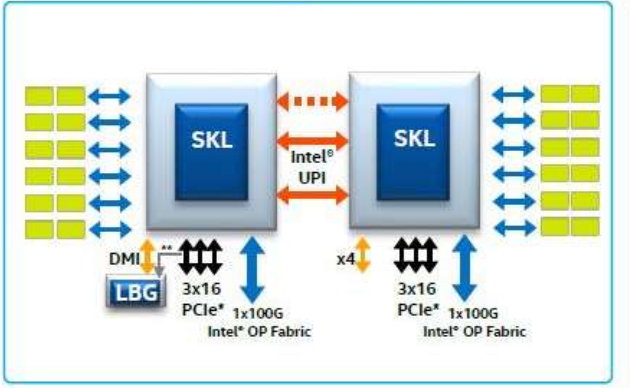
*

4.1 Overview of the Purley platform (6)

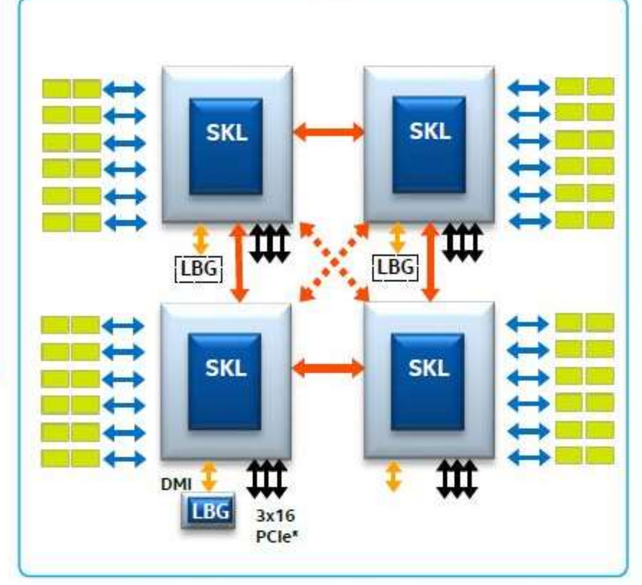
Reducing the diversity of the implementation of Intel's server lines by the Skylake-SP processor line in 2017:

It is achieved by unifying the processor implementations of the 2S, 4S and 8S configurations, as indicated in the following figure.

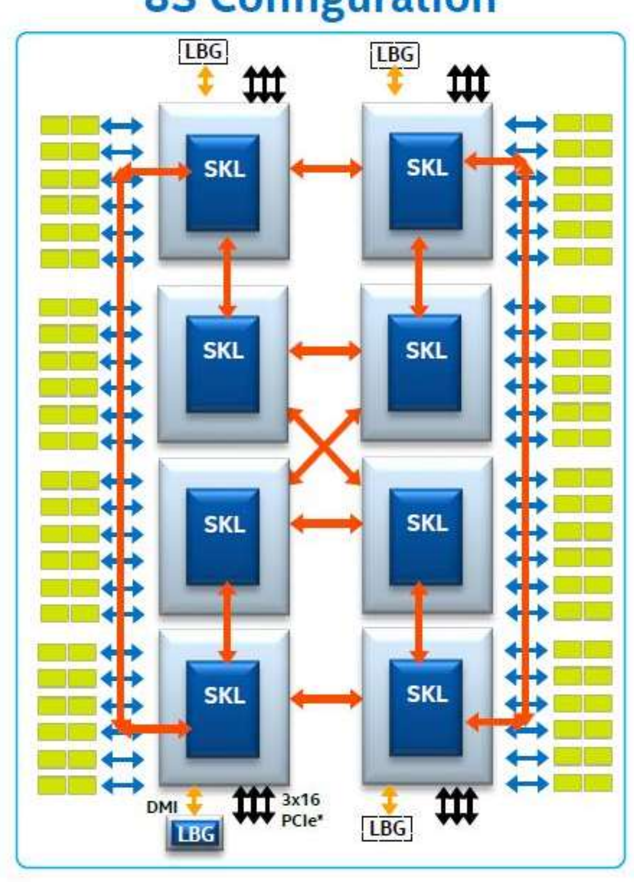
Unifying the processor implementations of the 2S, 4S and 8S configurations [139]



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



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



LBG: Lewisburg (PCH)
OP: OmniPath

Diversification of the existing performance categories -1

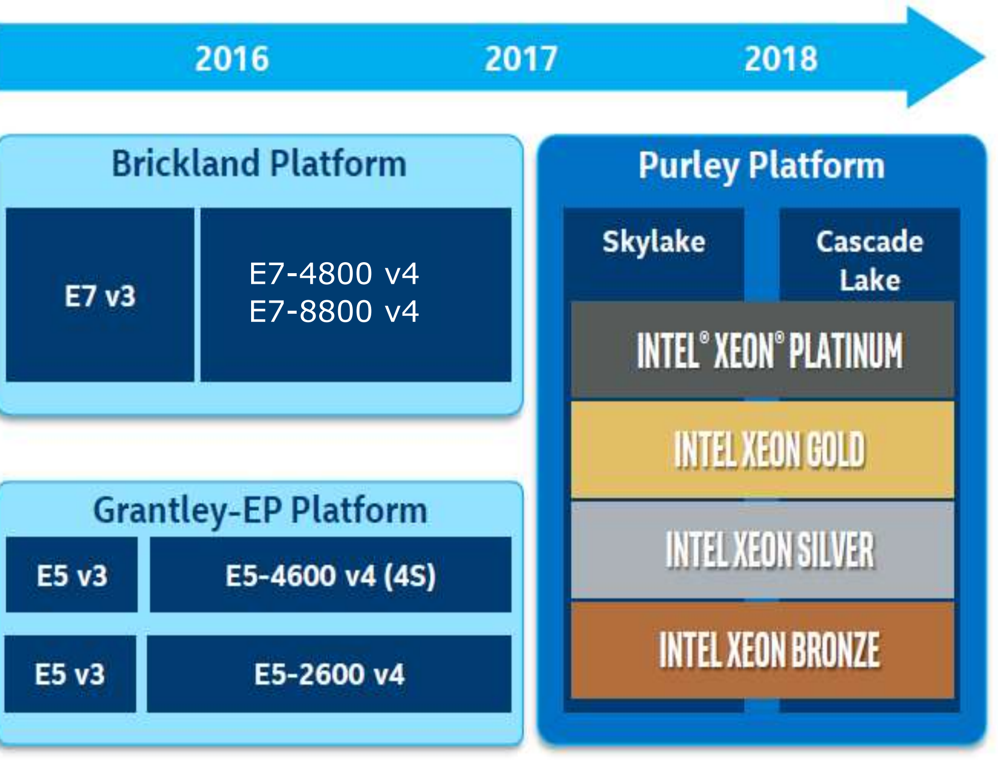
Instead of the existing E5 and E7 performance categories Intel introduced 4 different performance series, designated as follows:

- Platinum
- Gold
- Silver and
- Bronze

Diversification of the existing performance categories -2 [139]

Intel® Xeon® Processor E7
Targeted at **mission critical** applications that value a **scale-up** system with leadership **memory capacity** and **advanced RAS**

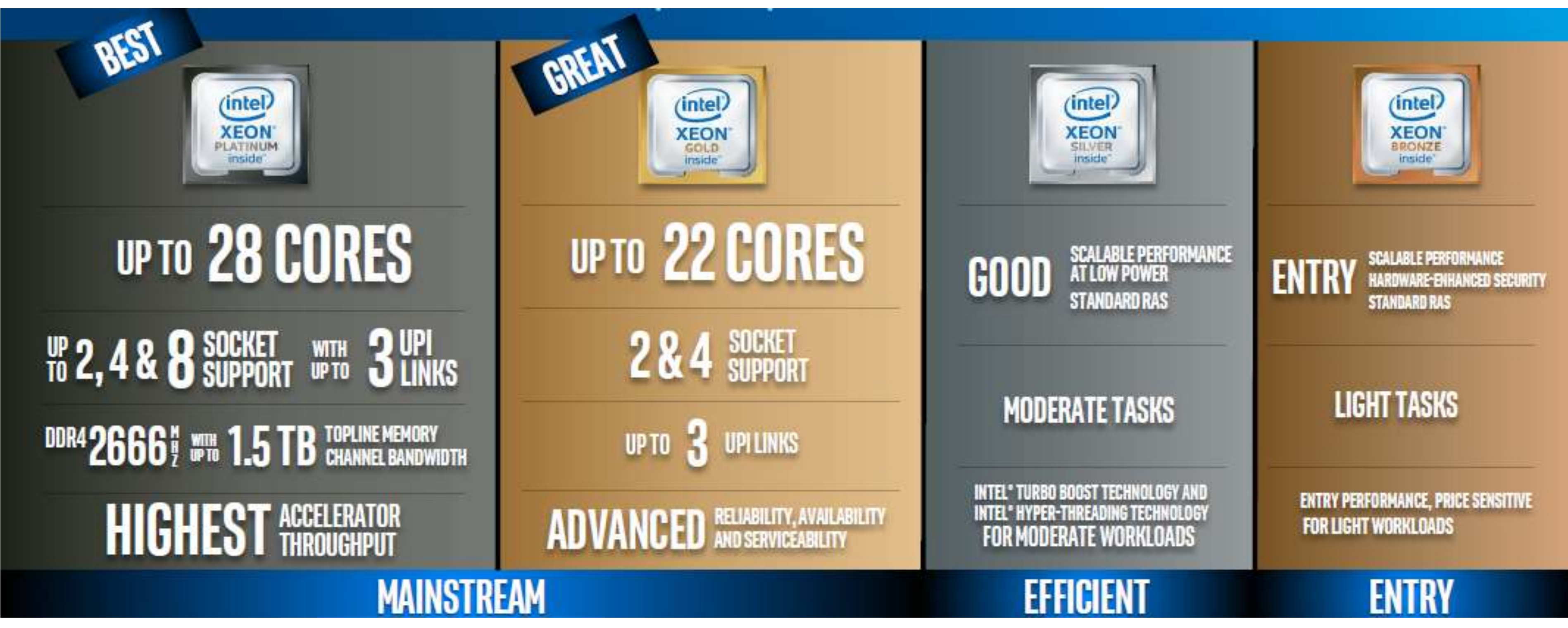
Intel® Xeon® Processor E5
Targeted at a wide variety of applications that value a **balanced system** with **leadership** performance/watt/$



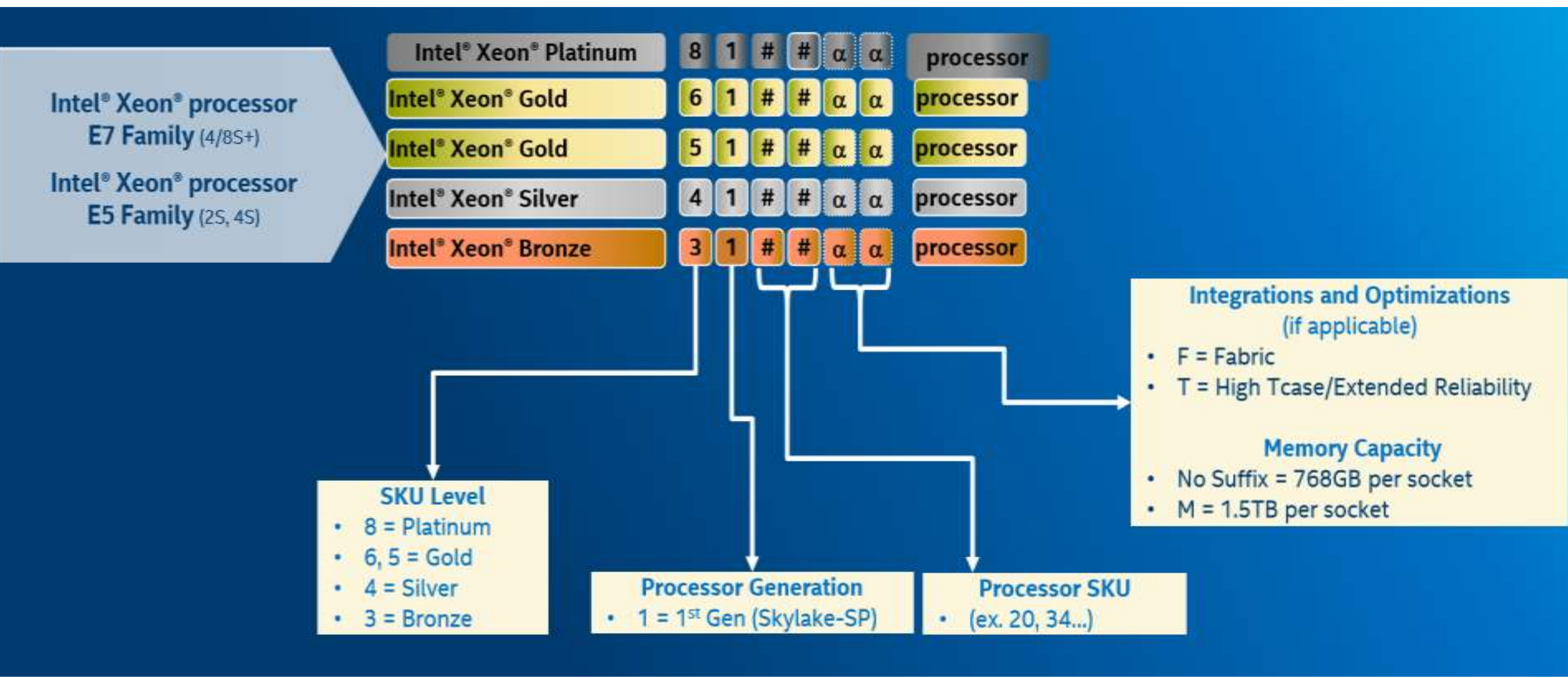
CONVERGED PLATFORM WITH INNOVATIVE SKYLAKE-SP MICROARCHITECTURE

*

Main features of the Platinum, Gold, Silver and Bronze series [143]

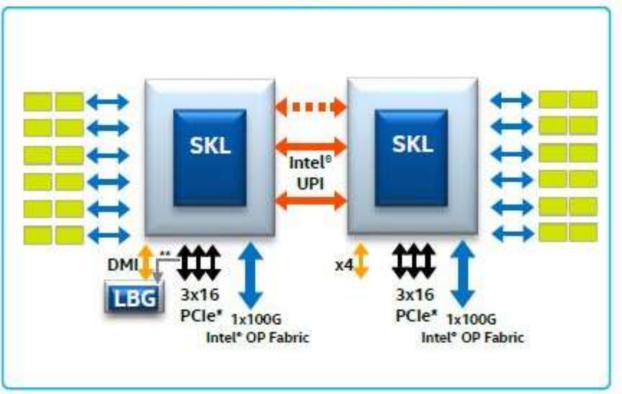


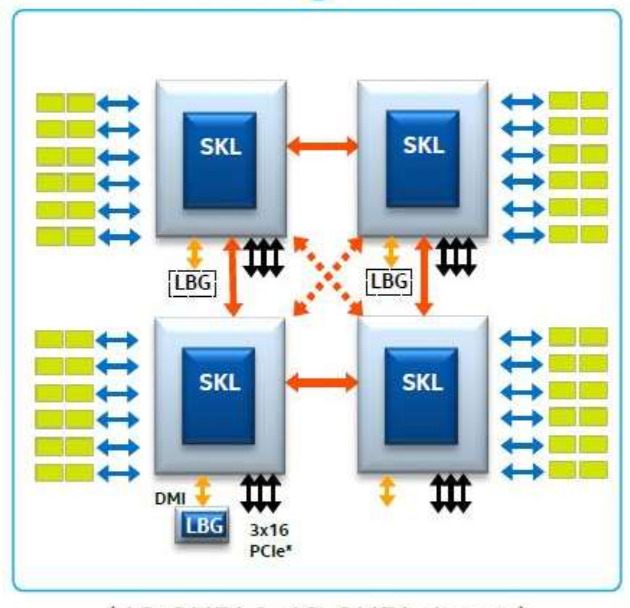
New model numbering for the Skylake-SP processor line [140]

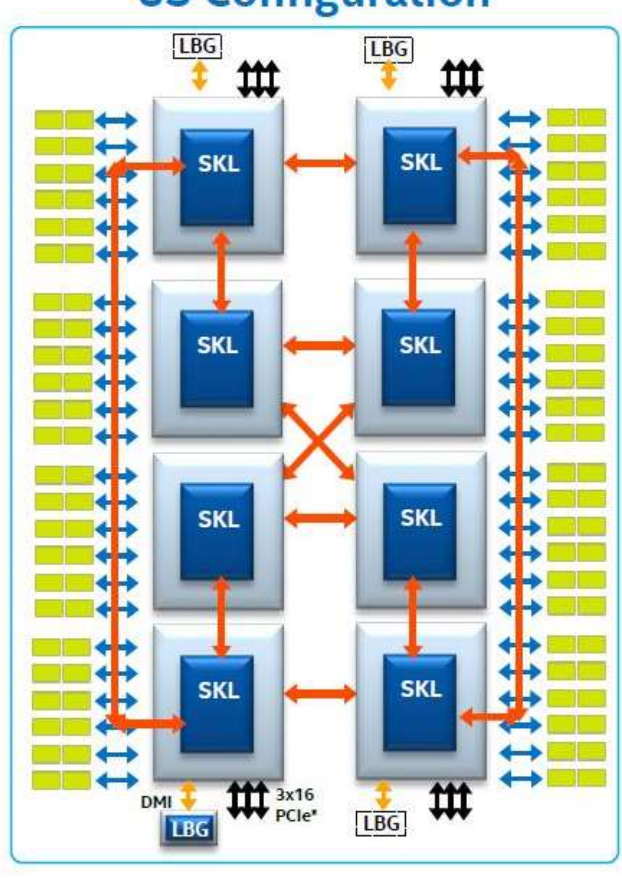


*

Supported platform topologies - ranging from 2S to 8S [139]

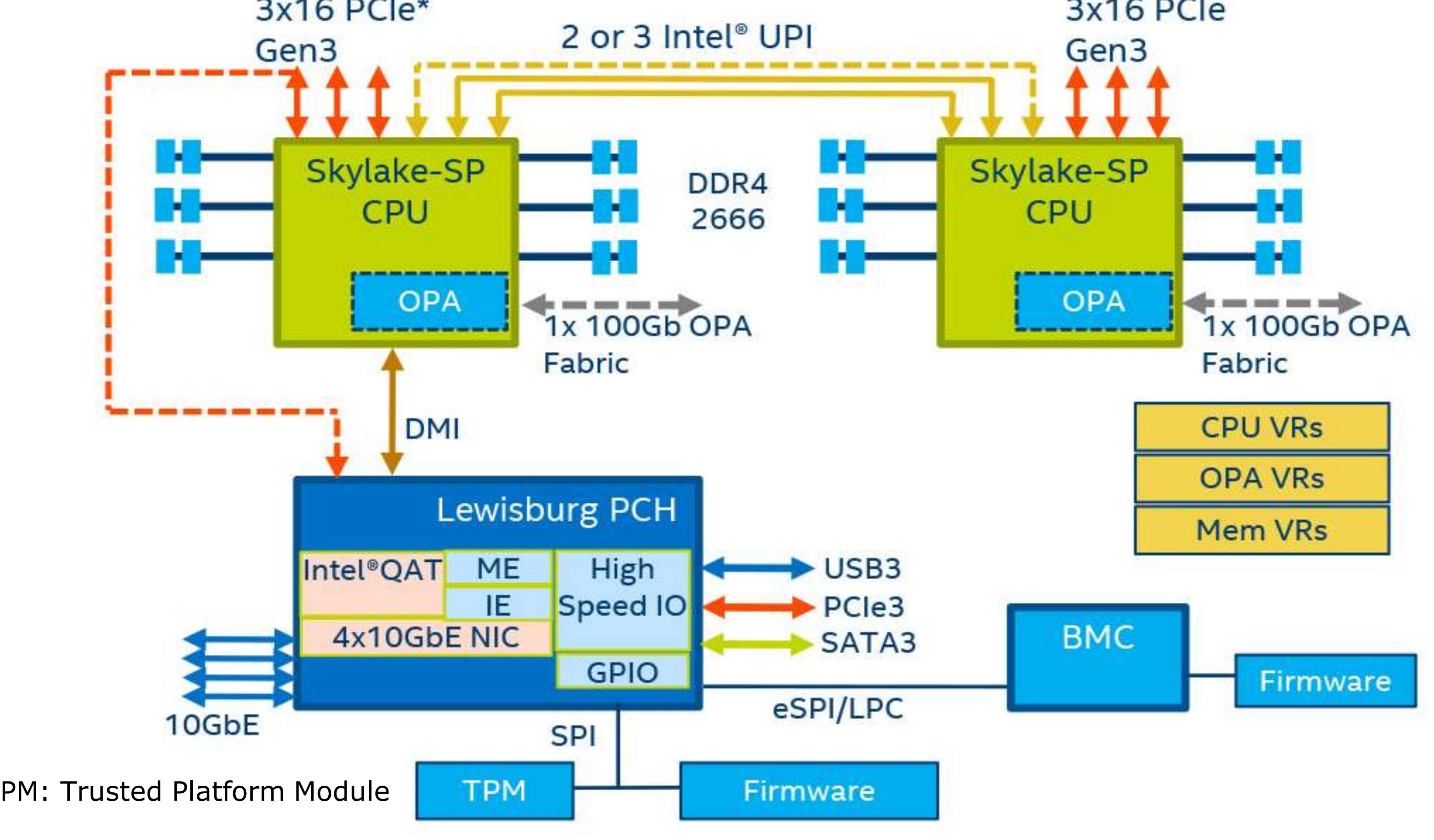






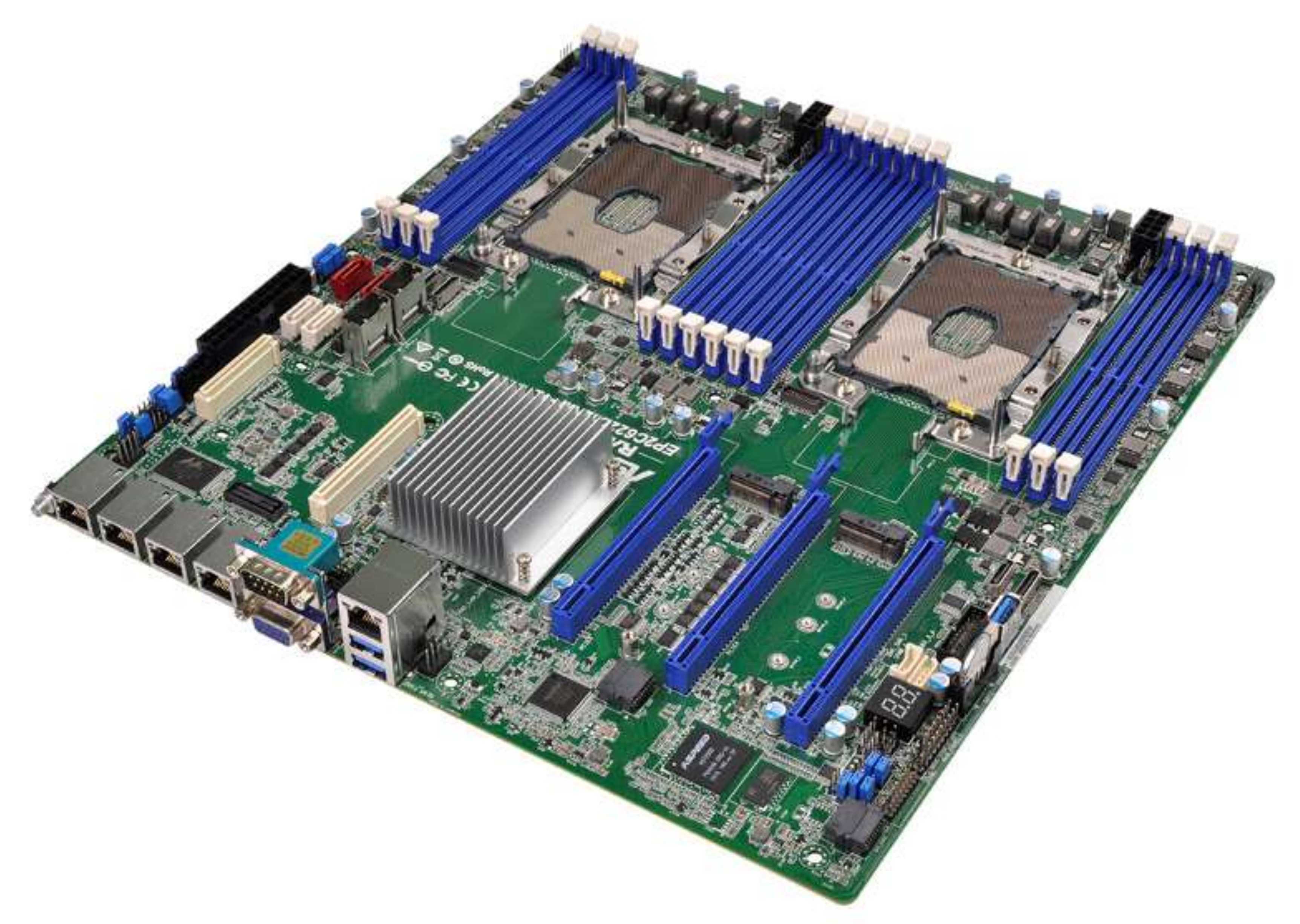
LBG: Lewisburg (PCH)
OP: OmniPath

Example: 2S server Skylake-SP configuration [139]

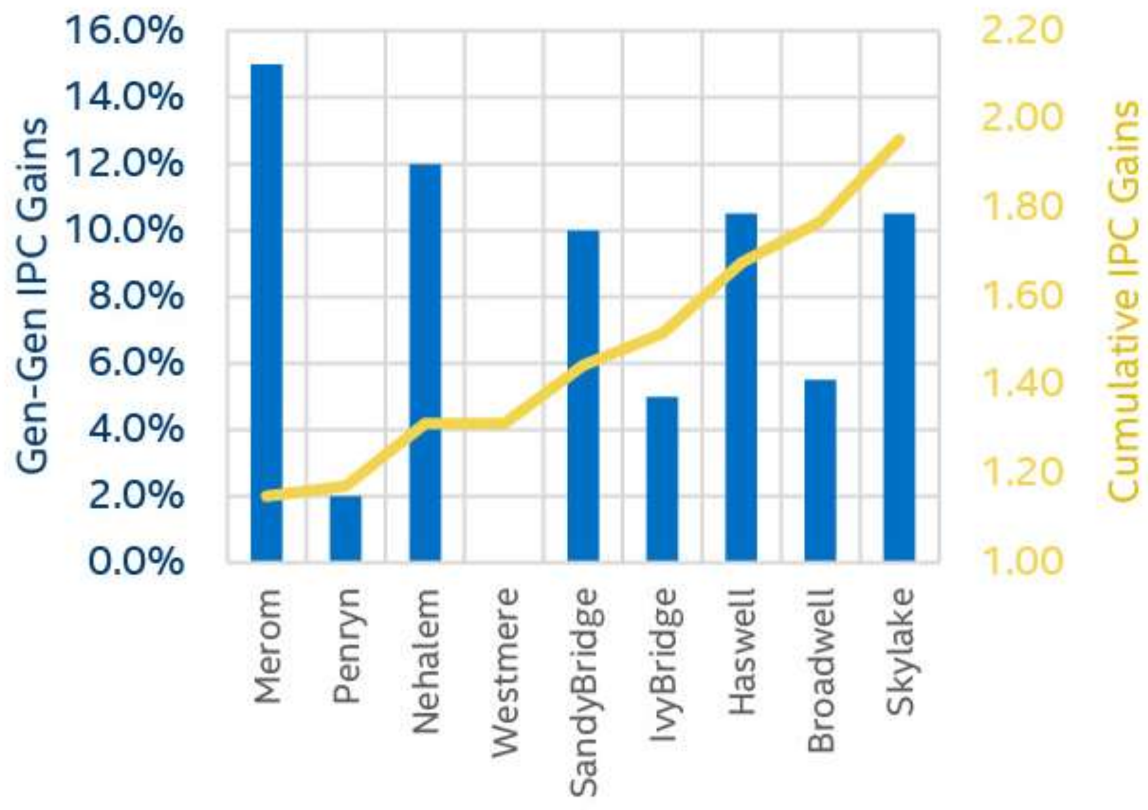


BMC: Baseboard Management Controller	PCH: Intel® Platform Controller Hub	IE: Innovation Engine
Intel® OPA: Intel® Omni-Path Architecture	Intel QAT: Intel® QuickAssist Technology	ME: Manageability Engine
NIC: Network Interface Controller	VMD: Volume Management Device	NTB: Non-Transparent Bridge

Example: Mainboard of a 2S server based on Skylake-SP processors [144]

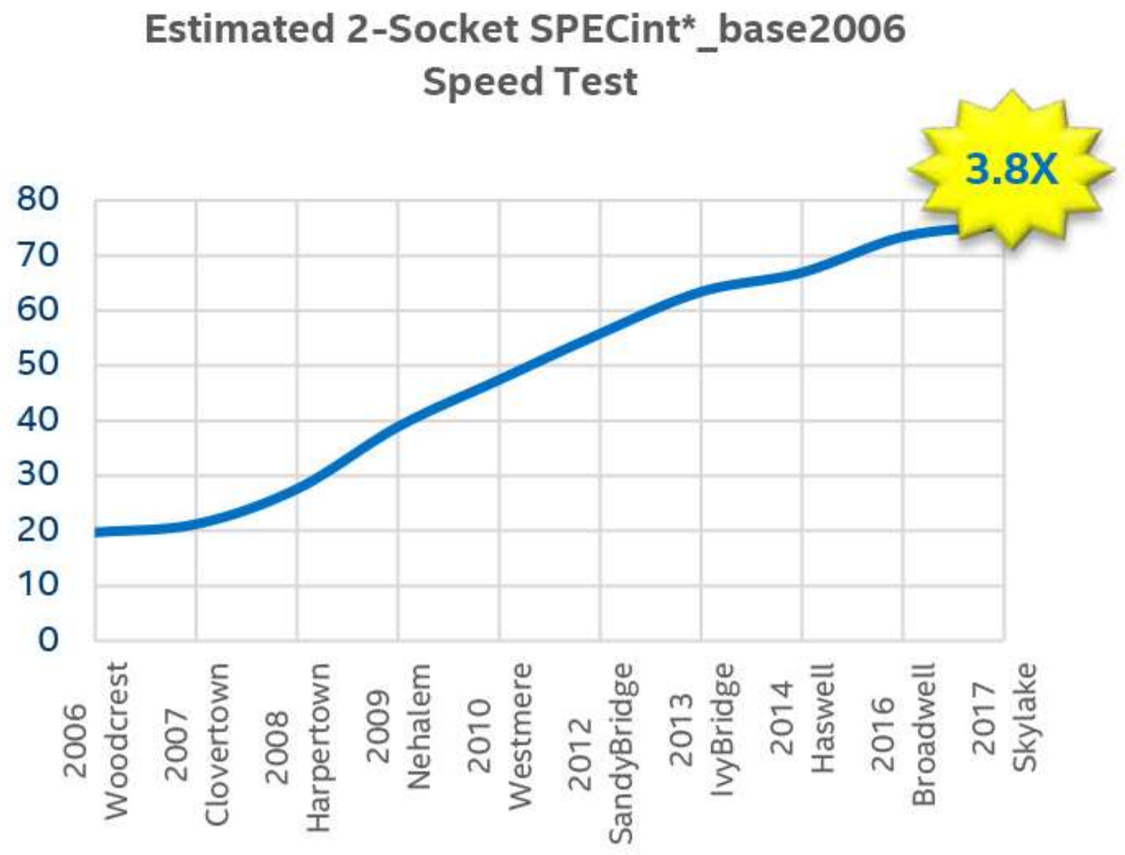


IPC improvement in Intel's Core 2 family (Core 2 to Skylake) [146]

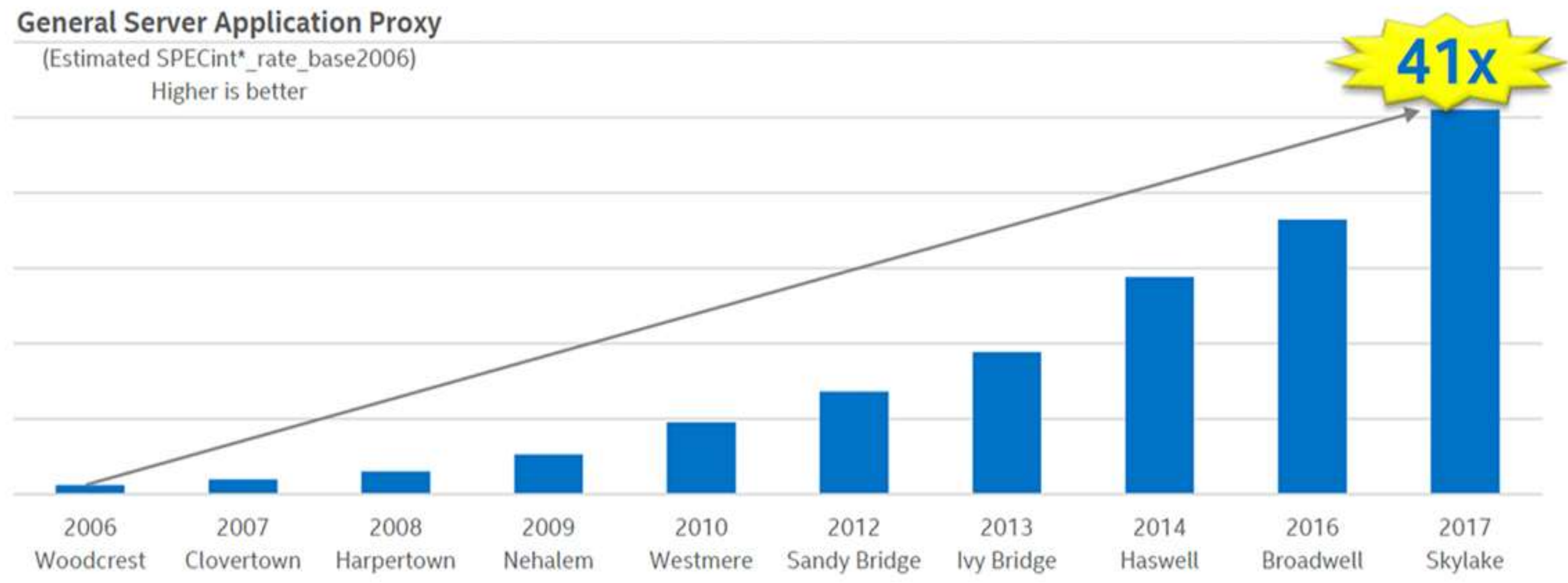


*

Single-thread performance improvement in Intel's 2S servers [146]

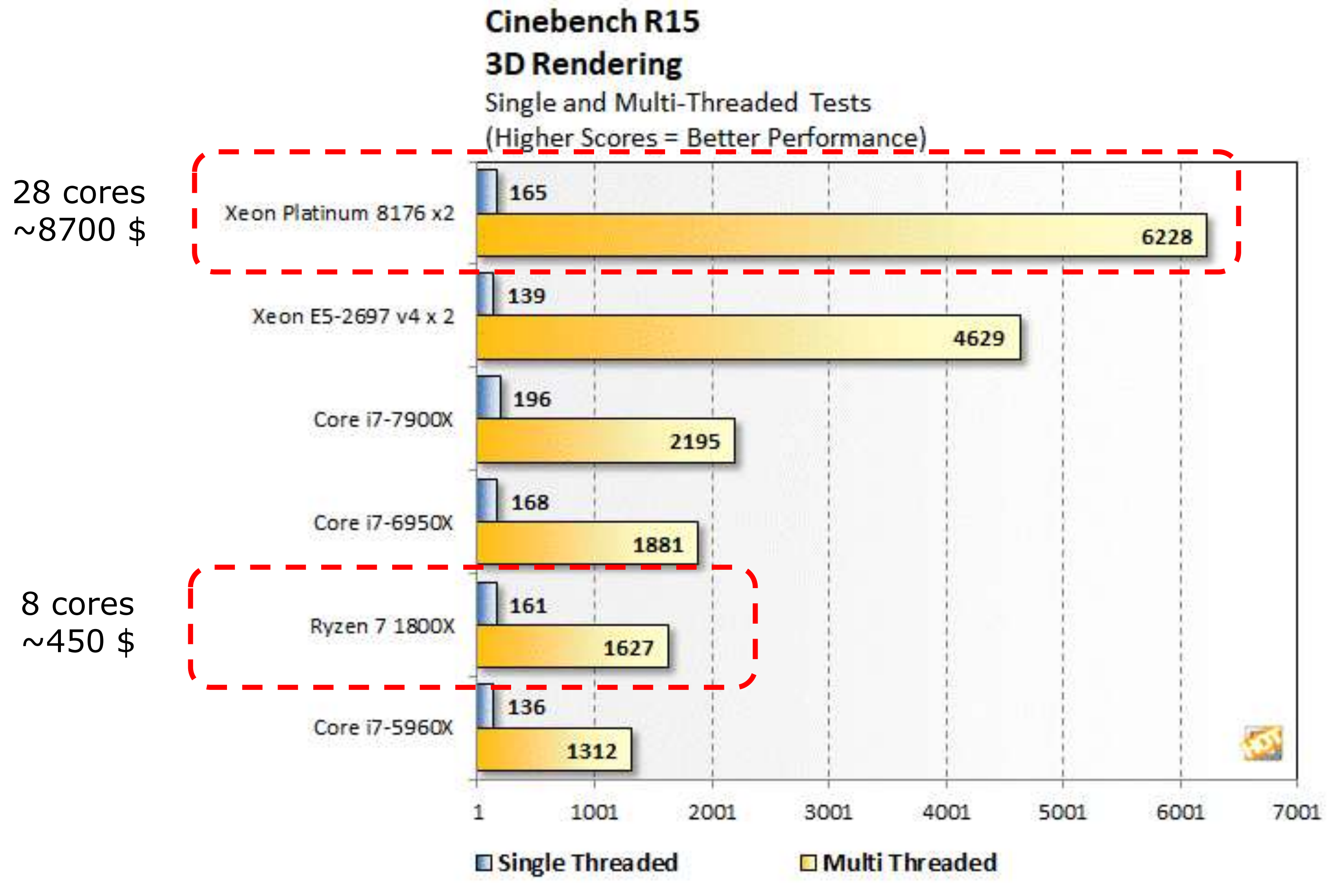


Multi-thread integer performance (throughput) improvement in Intel's 2S server lines [147]

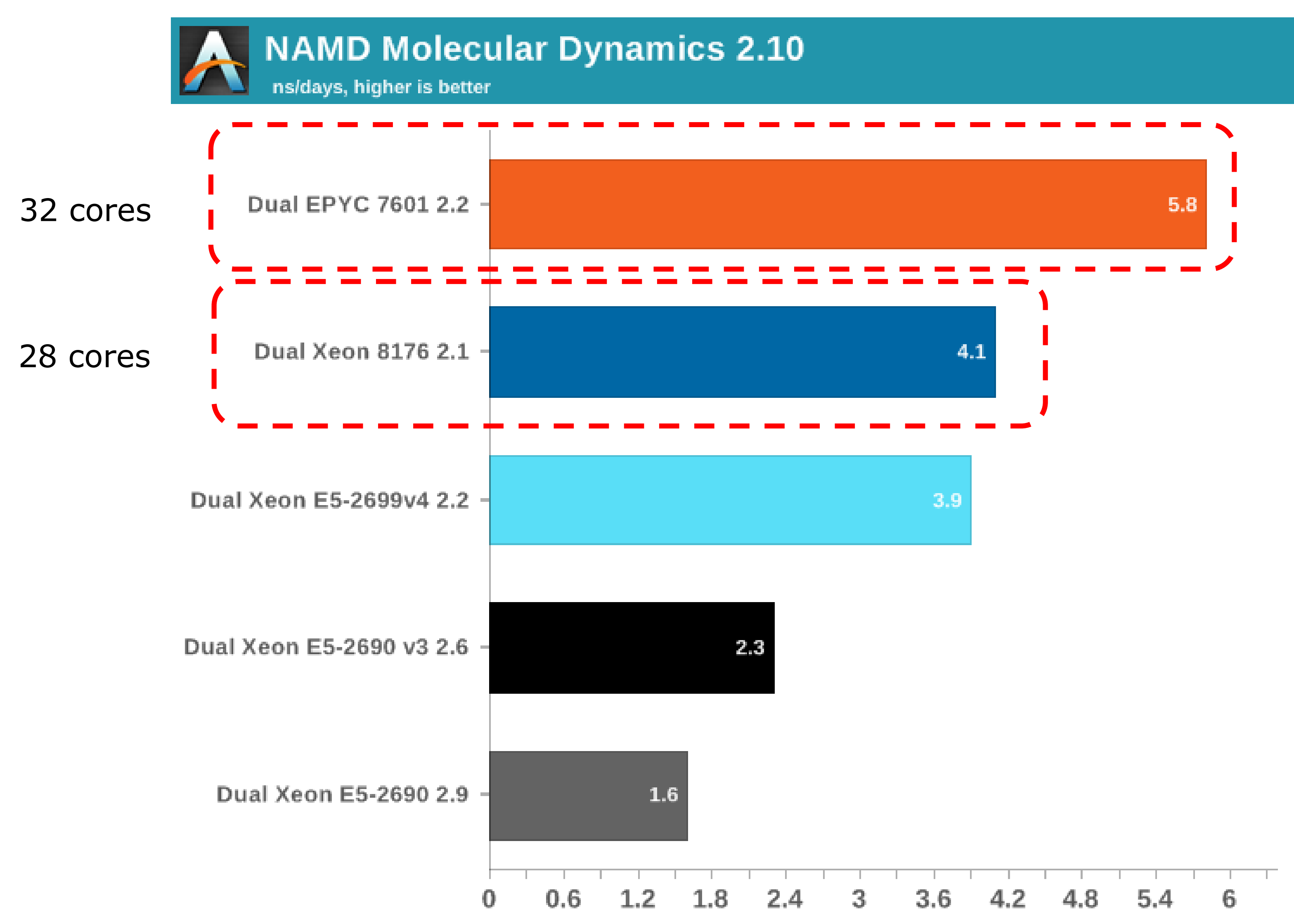


*

Benchmark results published by Intel for running the Cinebench R35 benchmark [147]



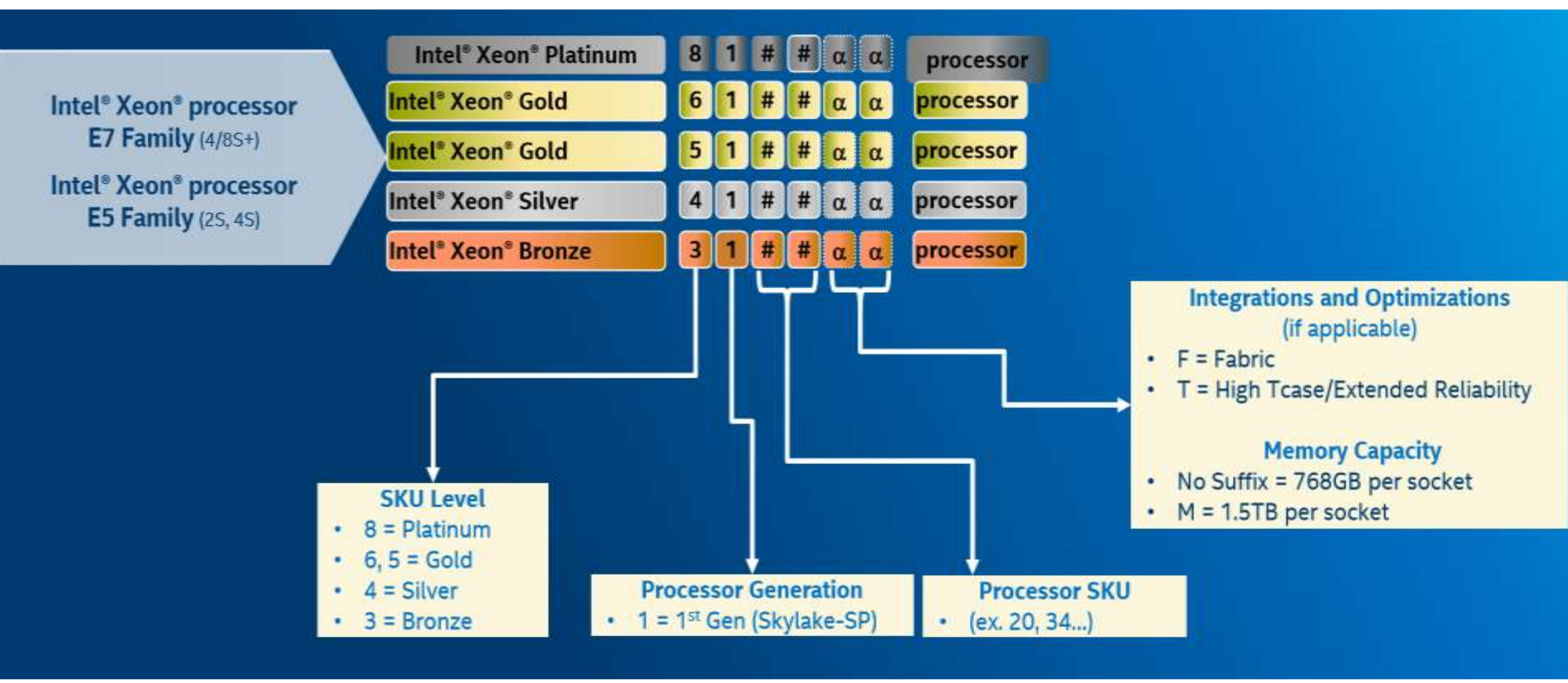
Benchmark result published by AMD for running molecular dynamics [148]



4.2 The Skylake-SP (Xeon x100) 2S/4S/8S processor line

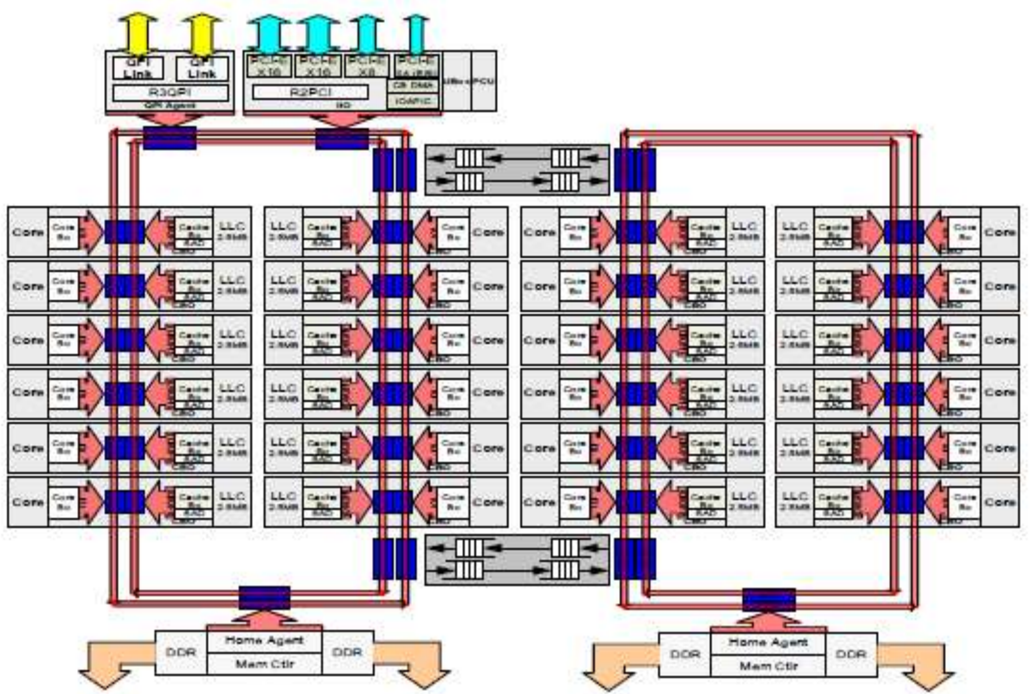
4.2.1 Overview

Model numbering for the Skylake-SP processor line [140]



*

Up to 28 cores [139]



Main features of the Skylake-SP Platinum, Gold, Silver and Bronze series [149]

	Intel® Xeon® Bronze Processor (3100 Series)	Intel® Xeon® Silver Processor (4100 Series)	Intel® Xeon® Gold Processor (5100 Series)	Intel® Xeon® Gold Processor (6100 Series)	Intel® Xeon® Platinum Processor (8100 Series)
PERVASIVE PERFORMANCE AND SECURITY					
Highest Core Count Supported	8 cores	12 cores	14 cores	22 cores	28 cores
Highest Supported Frequency	1.7 GHz (8C/85W)	2.2 GHz (10C/85W)	3.6 GHz (4C/105W)	3.4 GHz (6C/115W)	3.6 GHz (4C/105W)
Number of CPU Sockets Supported	Up to 2	Up to 2	Up to 4	Up to 4	Up to 8
Intel® Ultra Path Interconnect (UPI)	2	2	2	3	3
Intel® UPI Speed	9.6 GT/s	9.6 GT/s	10.4 GT/s	10.4 GT/s	10.4 GT/s
Intel® Advanced Vector Extensions 512 (AVX-512)	1 FMA	1 FMA	1 FMA	2 FMA	2 FMA
Memory Speed Support (DDR4)	2133 MHz	2400 MHz	2400 MHz	2666 MHz	2666 MHz
Highest Memory Capacity Supported Per Socket	768 GB	768 GB	768 GB	768 GB, 1.5 TB	768 GB, 1.5 TB
Intel® Omni-Path Architecture (Discrete PCIe* card)	•	•	•	•	•
Intel® QuickAssist Technology (Integrated in chipset)	•	•	•	•	•
Intel® QuickAssist Technology (Discrete PCIe card)	•	•	•	•	•
Intel® Optane™ Technology-based SSDs (3D XPoint™)	•	•	•	•	•
Intel® SSD Data Center Family (3D NAND)	•	•	•	•	•
PCIe 3.0 (48 lanes)	•	•	•	•	•
Intel® QuickData Technology (CBDMA)	•	•	•	•	•
Non-Transparent Bridge (NTB)	•	•	•	•	•
Intel® Turbo Boost Technology 2.0		•	•	•	•
Intel® Hyper-Threading Technology		•	•	•	•
Node Controller Support				•	•
Intel® Omni-Path Architecture (Integrated)				•	•
HIGH RELIABILITY					

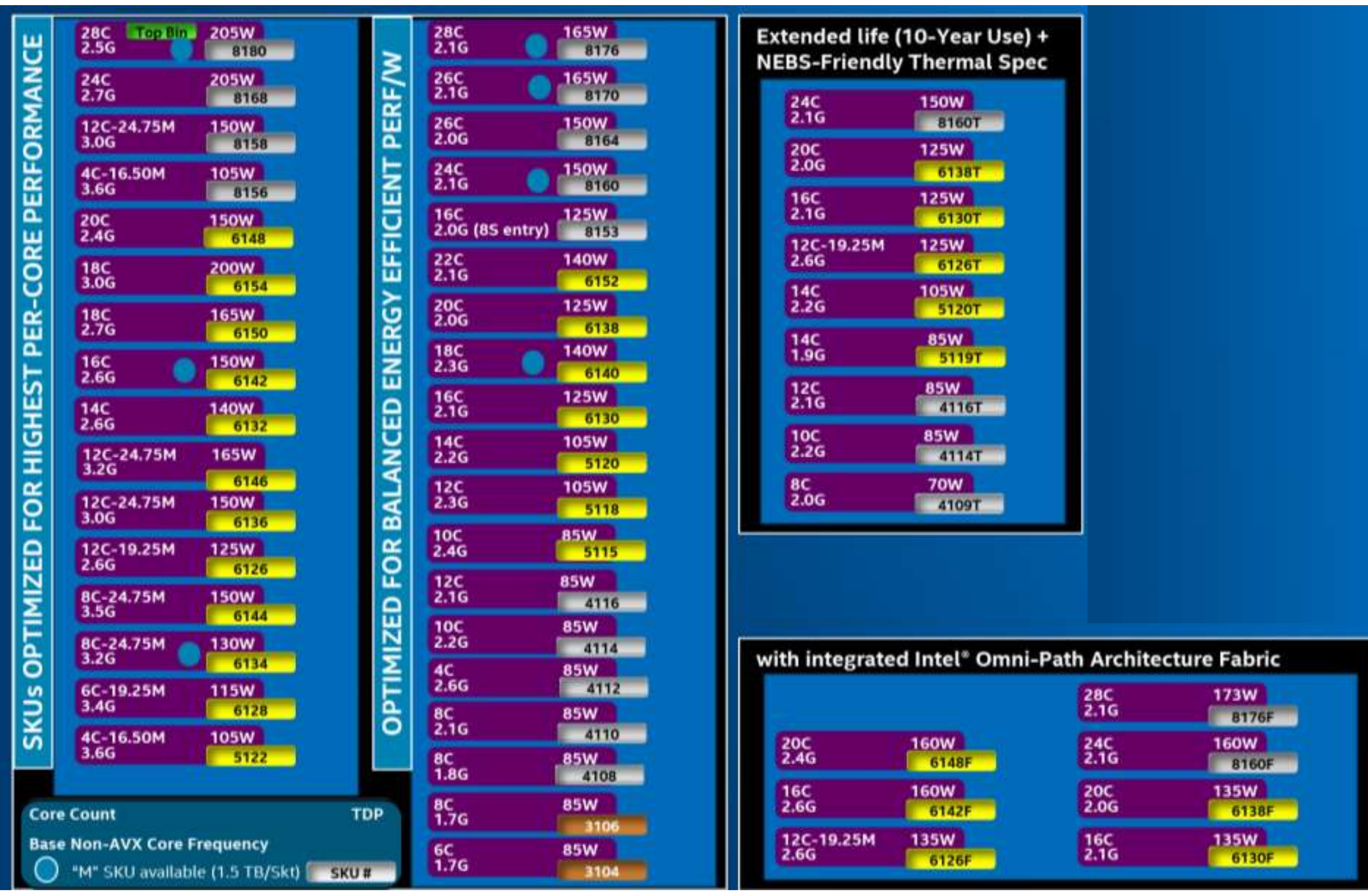
Up to four lineups in all four series (Platinum, Gold, Silver, Bronze) [146]

- Highly flexible offerings optimized to address broadest array of workloads
- Four primary SKU lineups:
  1. Optimized for per core performance
  2. Balanced, energy efficient perf/W
  3. Extended life (NEBS compliant)
  4. Integrated Intel® Omni-Path
- New 205W SKUs:
  - e.g. 8180 hits 2.5 GHz at 28 cores
- Great high frequency options:
  - e.g. 6144 w/ 8 cores at 3.5 GHz base
  - SKUs w/ add'l cache per core vs. default

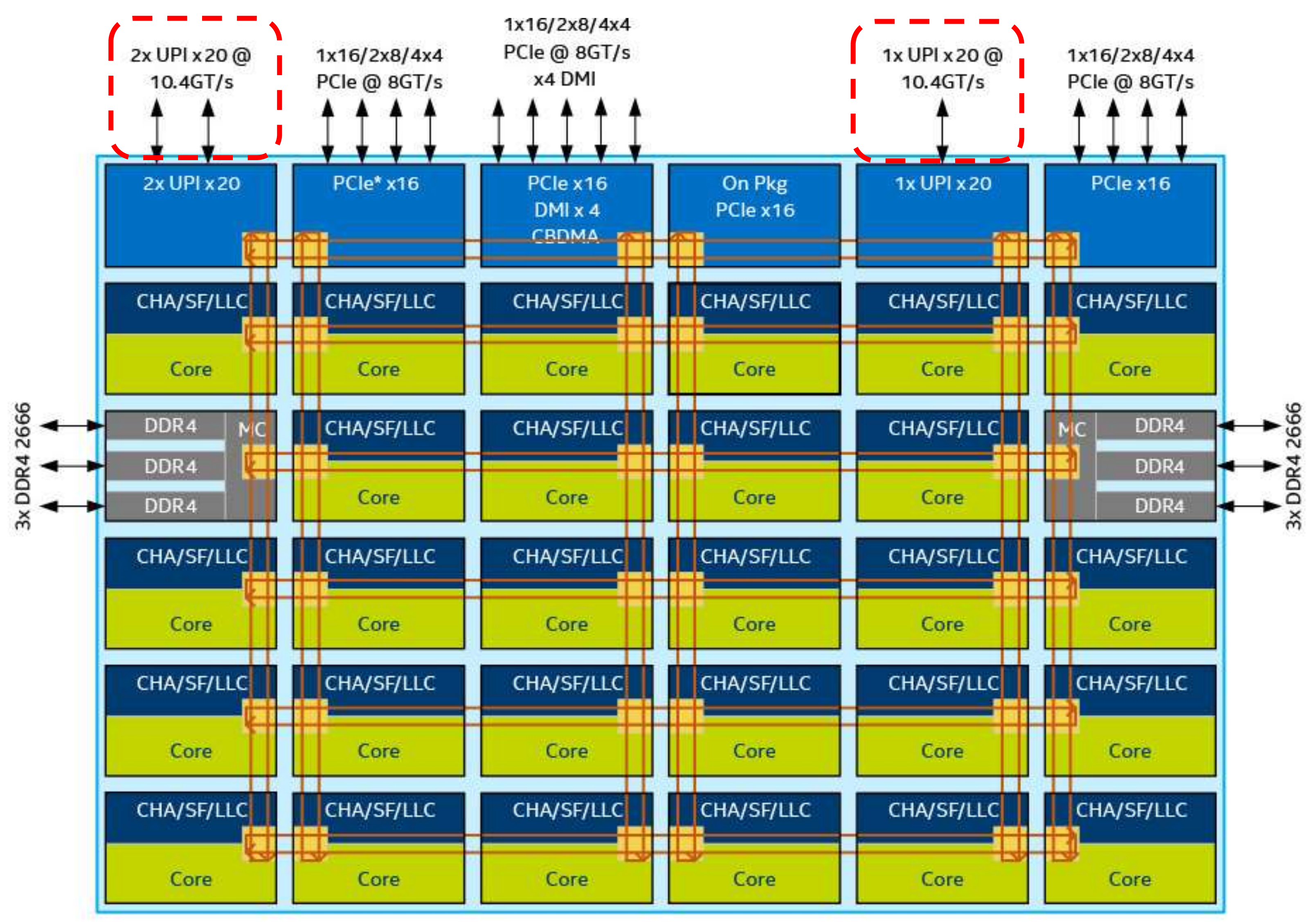
NEBS compliant [150]:

- NEBS (Network Equipment Building System) is a standard originating from Bell.
- It consists of various engineering requirements deemed essential to the reliability, availability and durability of equipment.

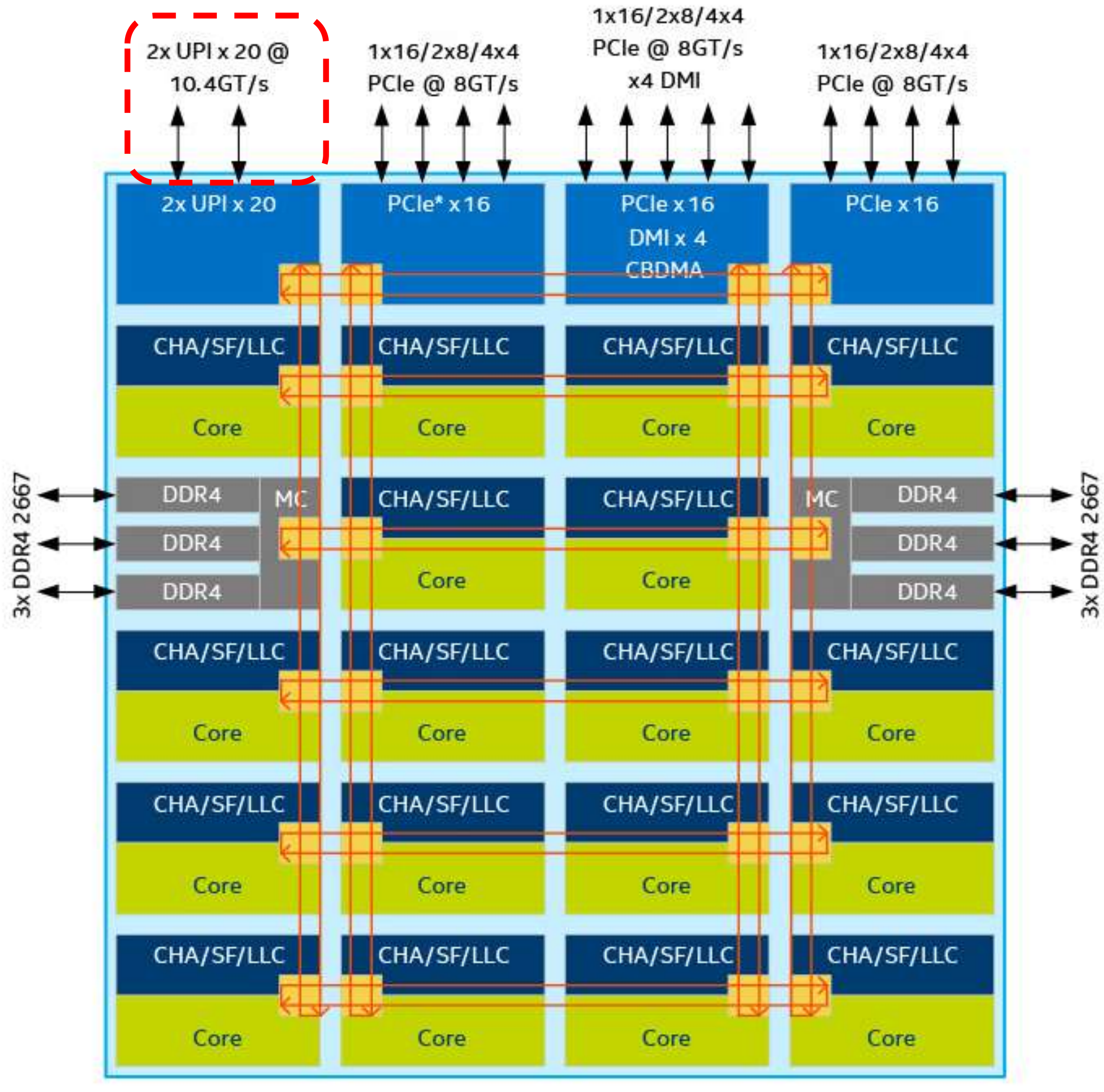
Main features of the introduced Skylake-SP models [146]



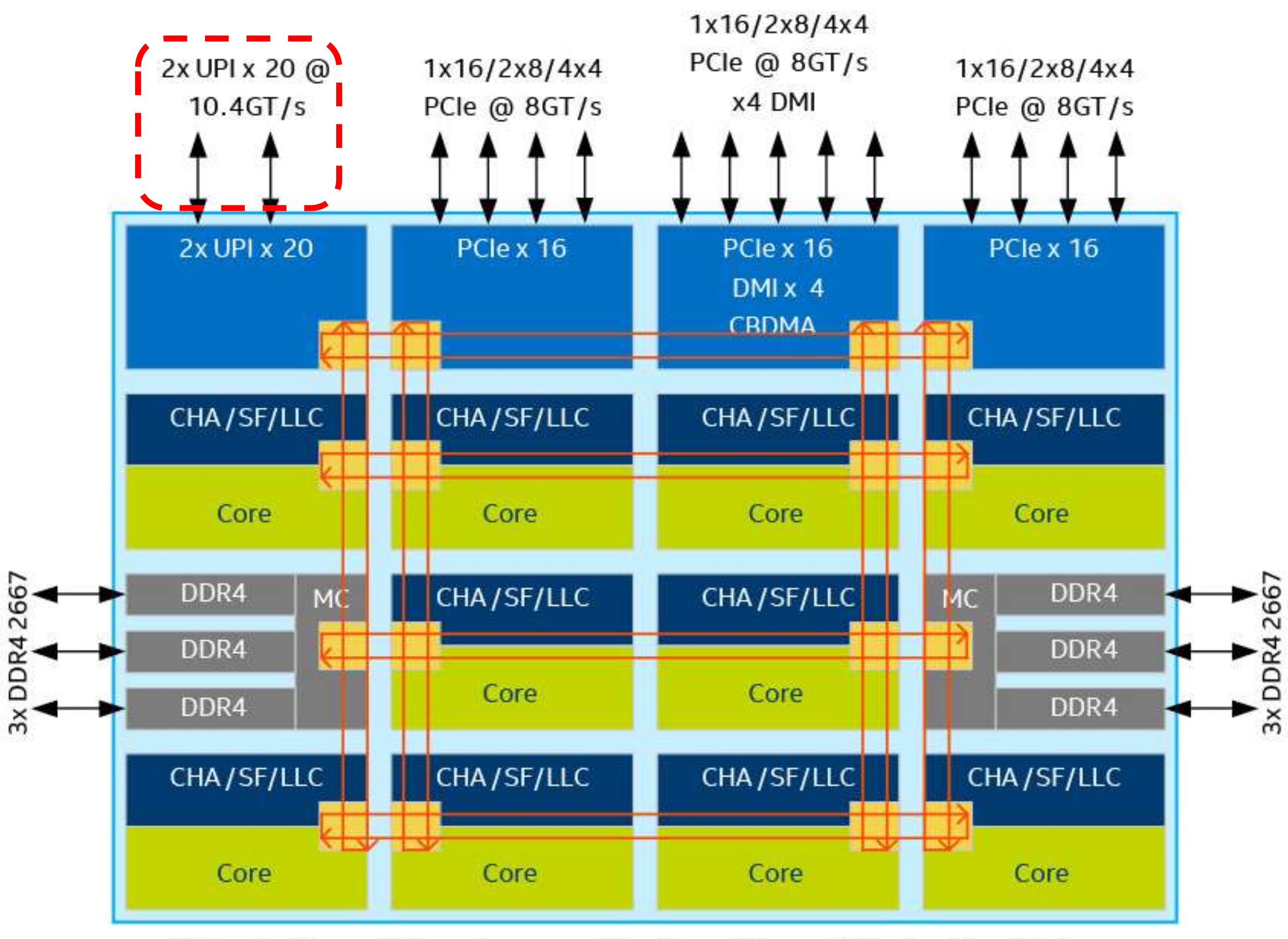
Die configurations - XCC die (up to 28 cores) [139]



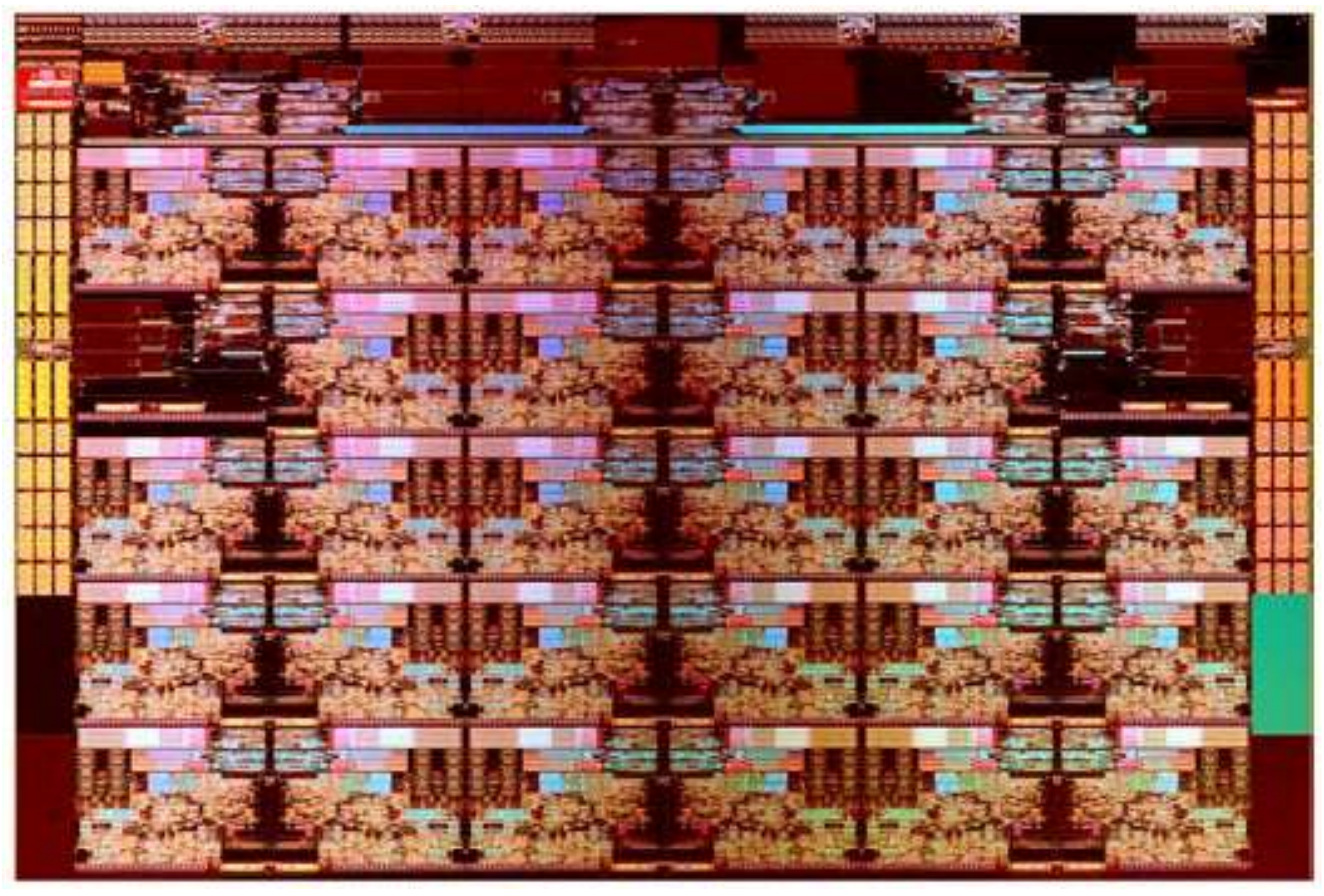
Die configurations - HCC die (up to 18 cores) [139]



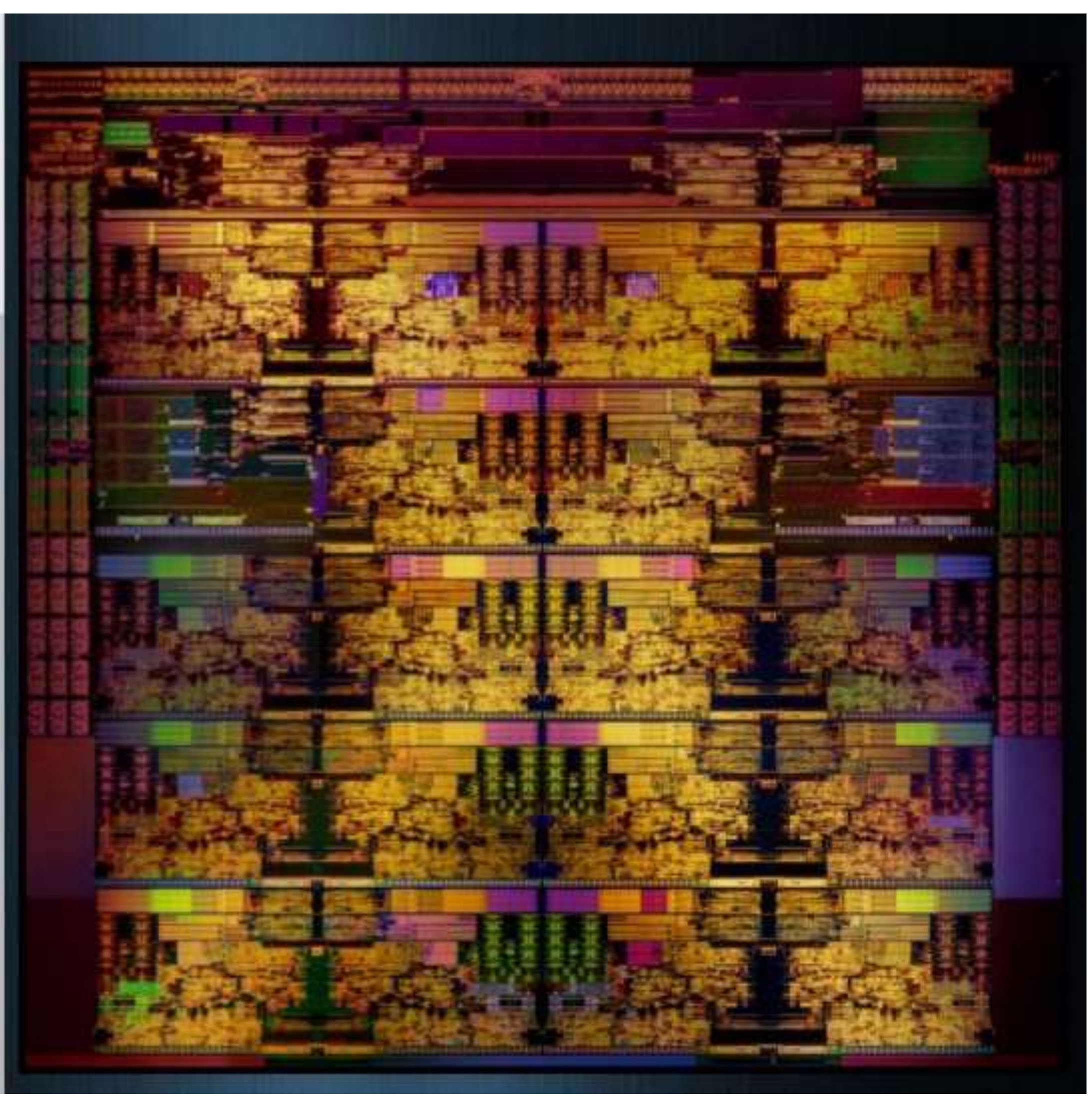
Die configurations - LCC die (up to 10 cores) [139]



The Skylake-SP XCC die (28 cores) [151]



The Skylake-SP HCC die (18 cores) [152]



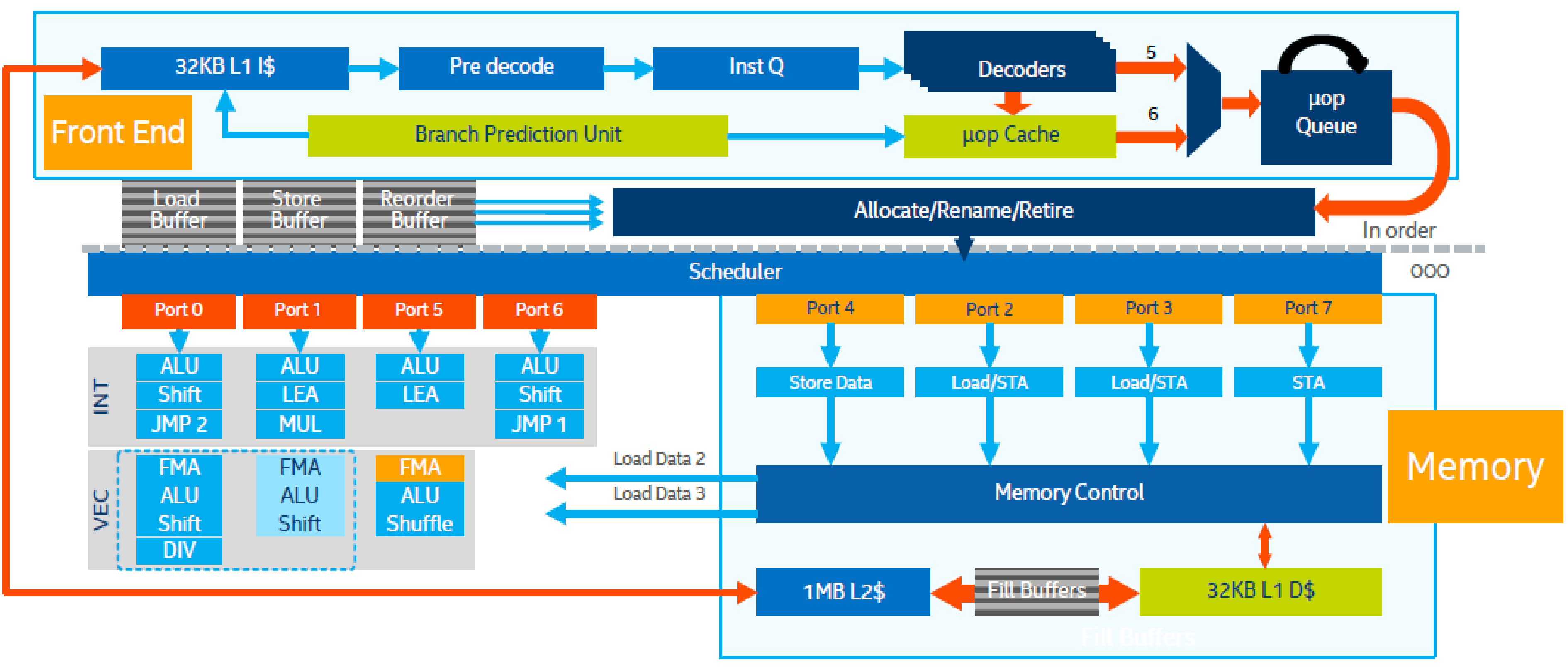
Expected die sizes [155]

Layout	Arrangement	Dimensions (mm)	Die Area (mm2)
LCC	3x4 (10-core)	22.0 x 14.0	308 mm^2
HCC	4x5 (18-core)	22.0 x 21.5	473 mm^2
XCC	5x6 (28-core)	21.5 x 31.5	677 mm^2

Note

Intel does not give any more die are figures, above data are estimates from [155].

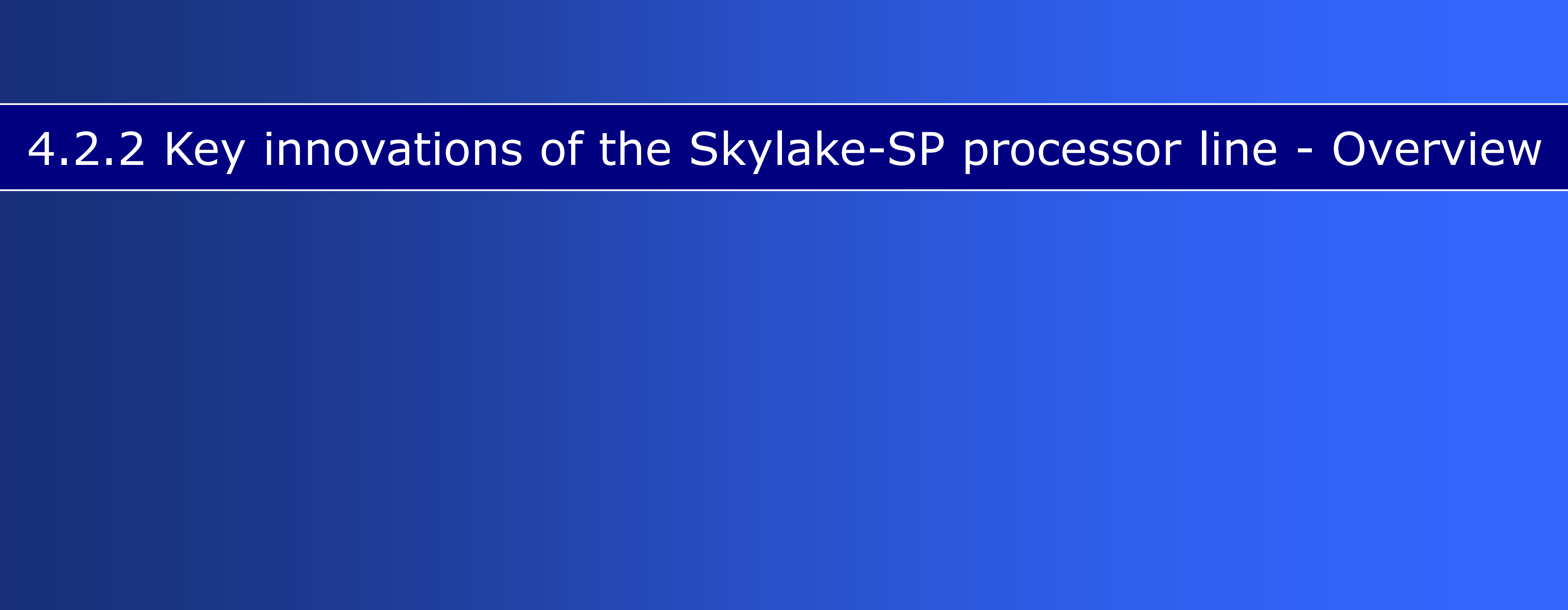
High-level overview of the microarchitecture of a Skylake-SP core [139]



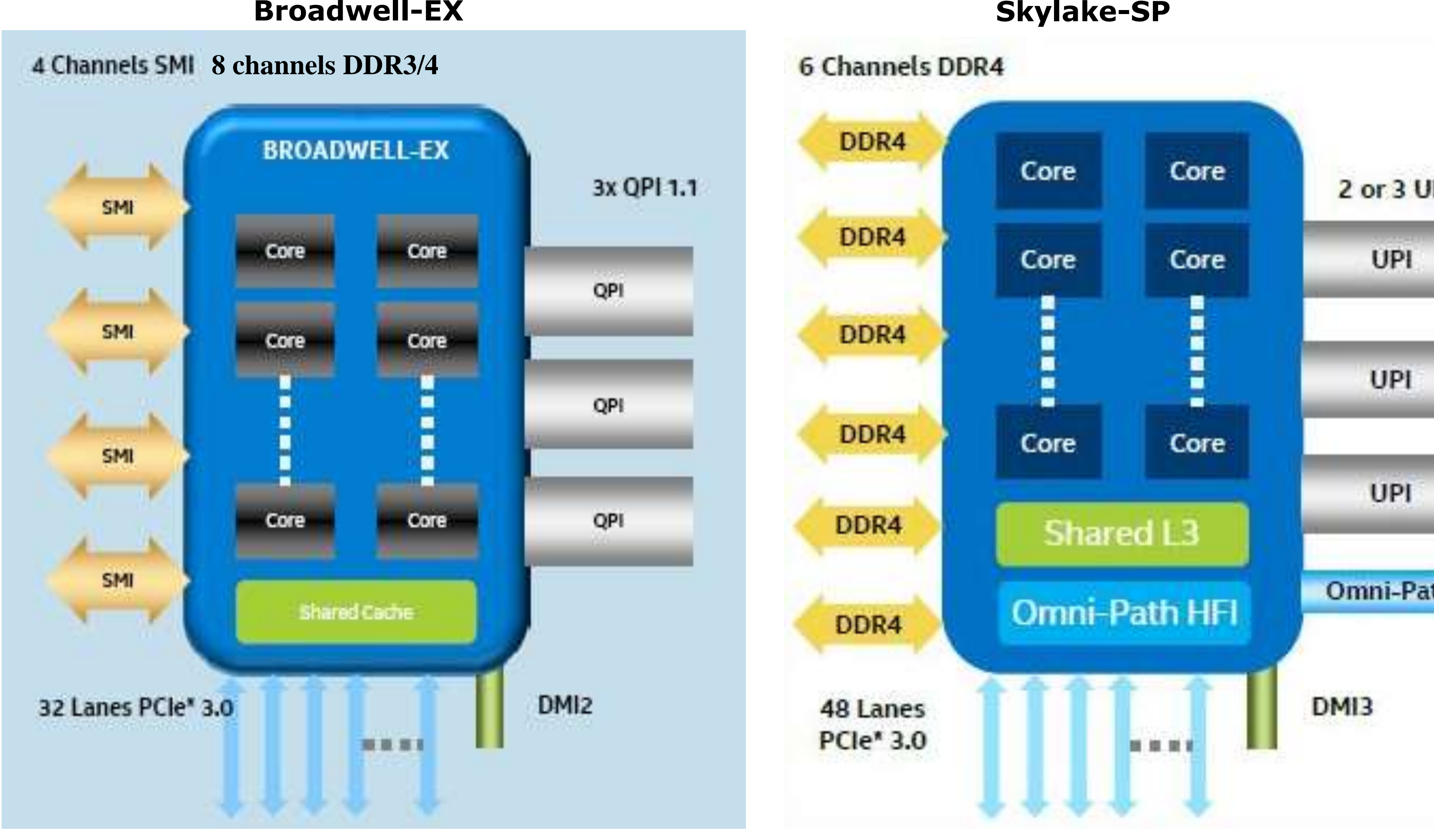
Quantitative improvements of the Skylake-SP microarchitecture vs. the previous one [139

	Broadwell uArch	Skylake uArch
Out-of-order Window	192	**224**
In-flight Loads + Stores	72 + 42	72 + **56**
Scheduler Entries	60	**97**
Registers – Integer + FP	168 + 168	**180** + 168
Allocation Queue	56	**64/thread**
L1D BW (B/Cyc) – Load + Store	64 + 32	**128 + 64**
L2 Unified TLB	4K+2M: 1024	4K+2M: **1536** **1G: 16**

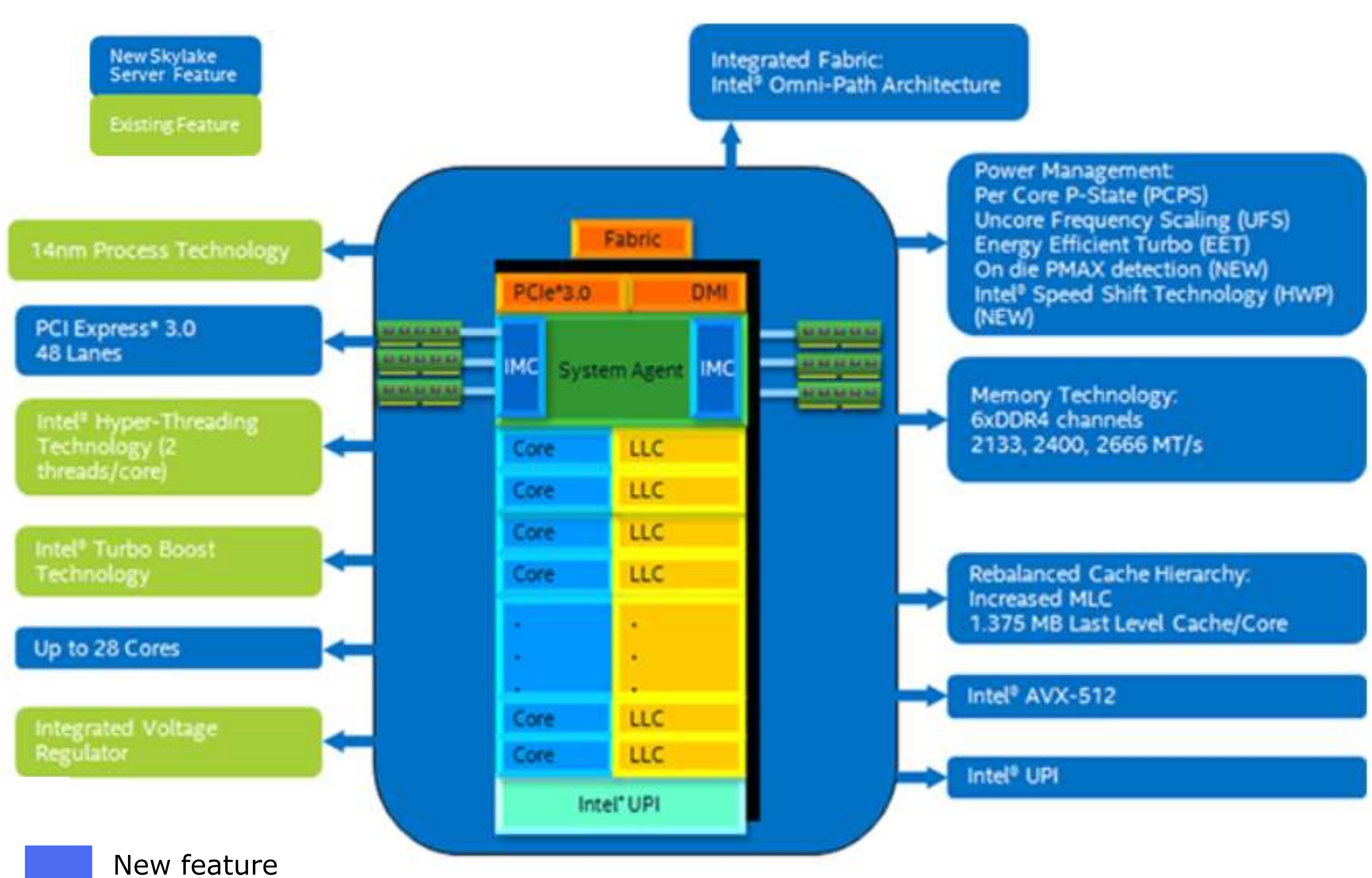
Out-of-order Window: ROB
TLB: Translation Lookaside Buffer


4.2.2 Key innovations of the Skylake-SP processor line - Overview

Contrasting key features of the Broadwell-EX and the Skylake-SP processors [153], [139]



New Skylake-SP features [154]



4.2.2 Key innovations of the Skylake-SP processor line - Overview (3)

Key innovations of the Skylake-SP processor line discussed

a. AVX512
b. Re-architected L2/L3 cache hierarchy
c. 6 direct attached DDR4 memory channels
d. Mesh architecture
e. UPI (Ultra Path Interconnect)
f. In-package integrated OmniPath host fabric interface
g. New socket

*

a) AVX512

a) AVX512

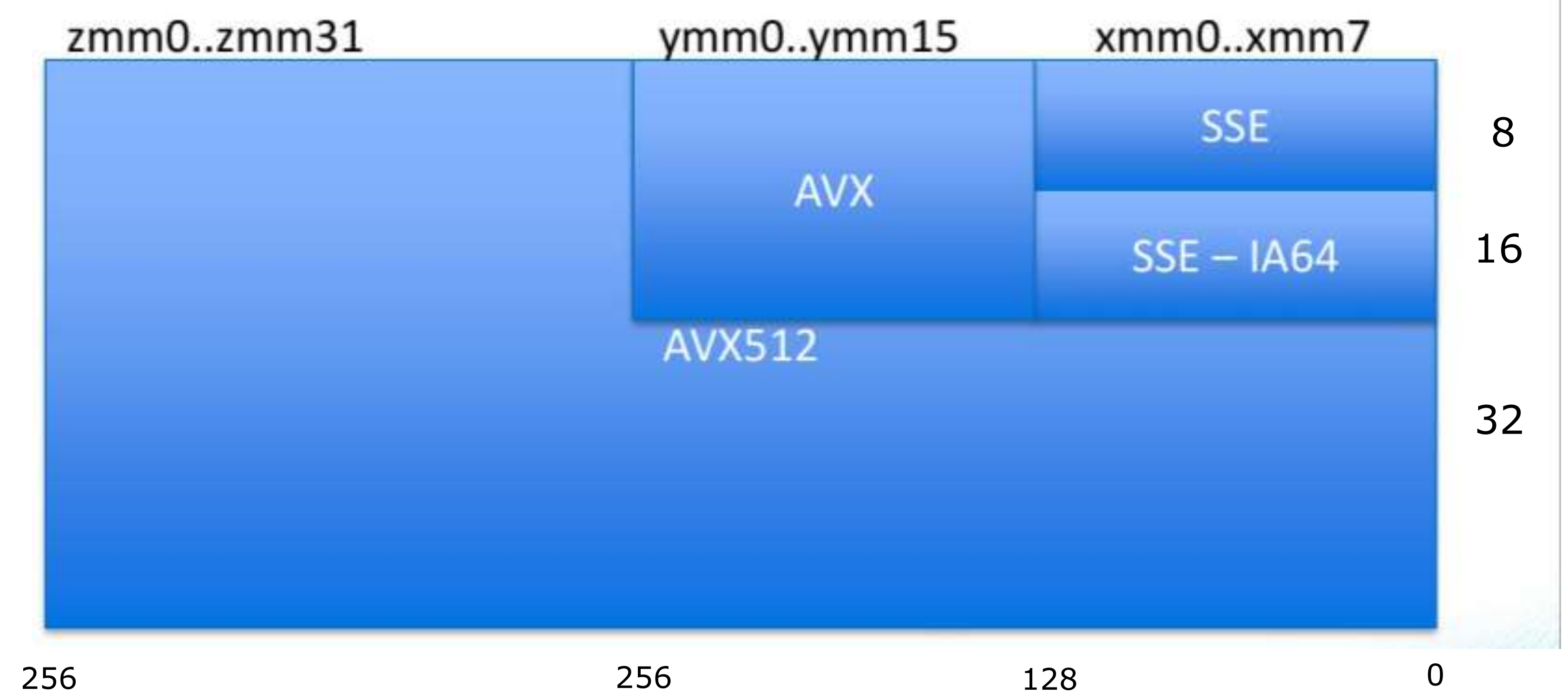
Evolution of Intel's SIMD extensions

Designation	Intro.	Processor line	Technology	SIMD registers	SIMD Register set	Instruction set
MMX	1997	Pentium MMX	350 nm	MM [0:7][1]	8x64 bit	FX SIMD
SSE	1999	Pentium III	250 nm	XMM [0:7]	8x128 bit	FX/FP SIMD
SSE2	2000	Pentium 4	180 nm	XMM [0:15]	16x128 bit	FX/FP SIMD
AVX	2011	Sandy Bridge	32 nm	YMM [0:15]	16x256 bit	FP SIMD
AVX2	2013	Haswell	22 nm	YMM [0:15]	16x256 bit	FX/FP SIMD
AVX512	2017	Skylake-SP	14 nm	ZMM [0:31]	32x512 bit	FX/FP SIMD

[1]The MM registers are aliased with the mantissa part of the FP registers

*

Extension of the available SIMD register space [155]

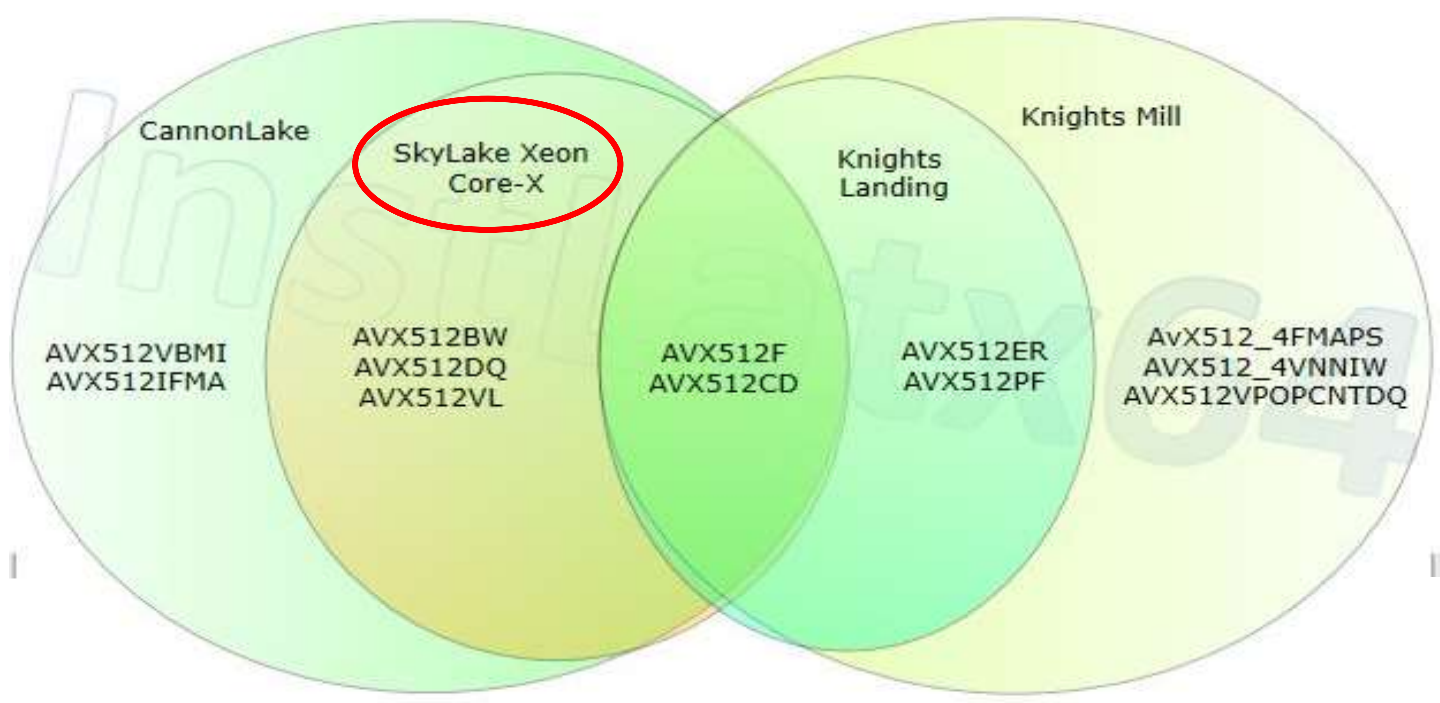


a) AVX512 (3)

Evolution of the SP/FP and DP/FP performance in Intel's Core 2 lines [139]

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

Different versions of the AVX512 instruction set [155] Source: Intel SDE 8.40 (2017-06-01)

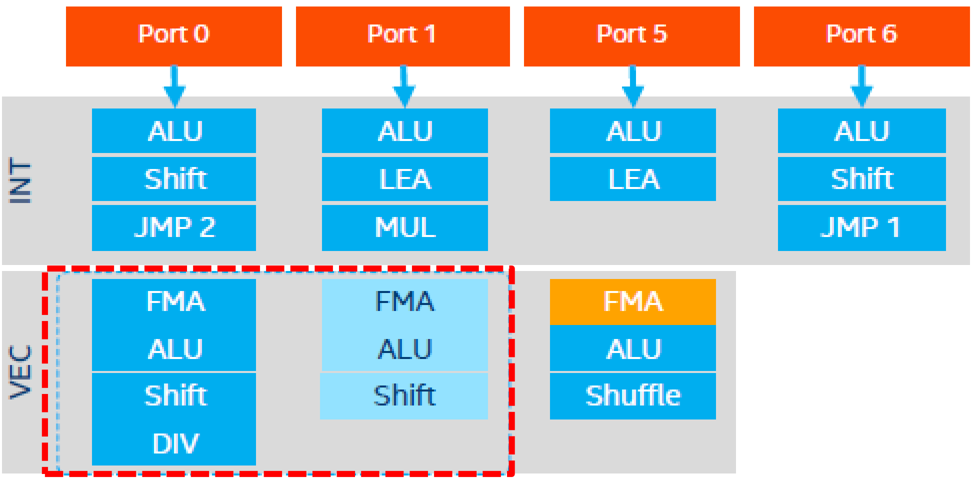


AVX-512-F: Foundation instructions
AVX-512-CD: Conflict Detect (loop vectorization with possible conflicts)
AVX-512-BW: Support for 512-bit Word support
AVX-512-DQ: More instructions for double/quad math operations
AVX-512-VL: Foundation plus <512-bit vector length support
AVX-512-ER: Exponential and Reciprocal
AVX-512-IFMA: Integer Fused Multiply Add with 52-bit precision
AVX-512-PF: Prefetch Instructions
AVX-512-VBMI: Vector Byte Manipulation Instructions
AVX-512-4VNNIW: Vector Neural Network Instructions Word (variable precision)
AVX-512-4FMAPS: Fused Multiply Accumulation Packed Single precision

Implementation of AVX512 in Skylake-SP [139]

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

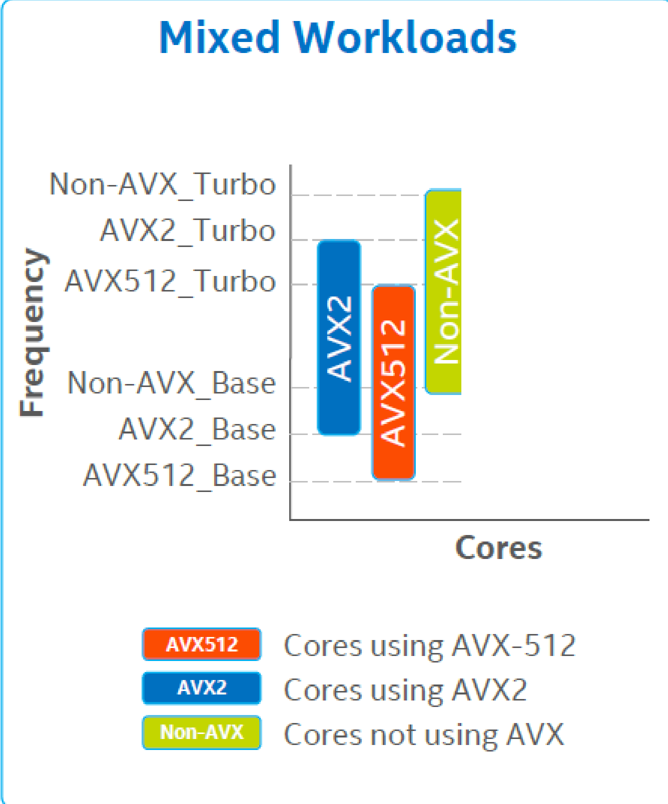


*

Reduced turbo frequencies while running AVX code [139]

- 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

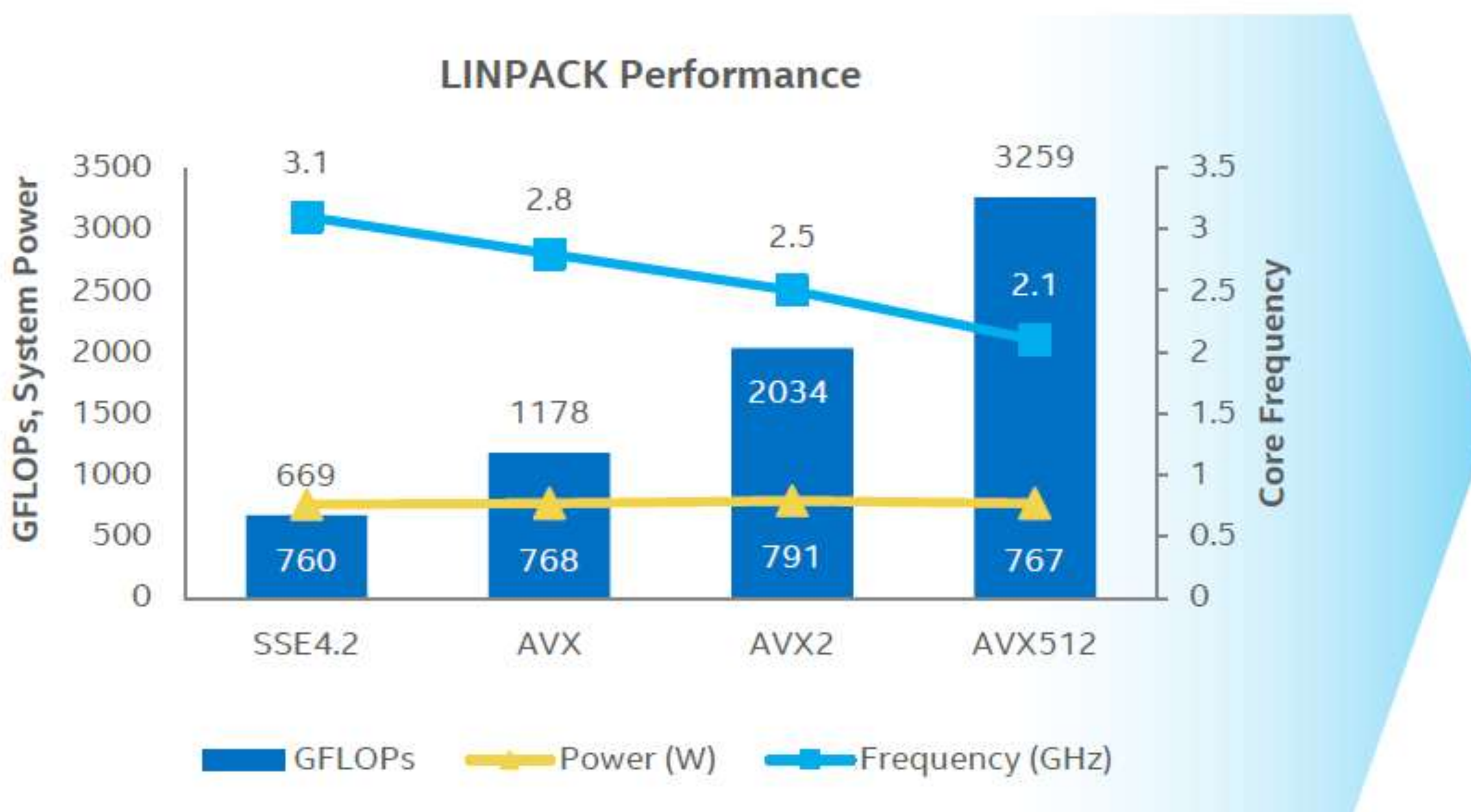
Code Type	All Core Frequency Limit
SSE AVX2-Light (without FP & int-mul)	Non-AVX All Core Turbo
AVX2-Heavy (FP & int-mul) AVX512-Light (without FP & int-mul)	AVX2 All Core Turbo
AVX512-Heavy (FP & int-mul)	AVX512 All Core Turbo



Note that turbo frequency limits are controlled on a per core basis.

*

Performance increase over SIMD generations [139]



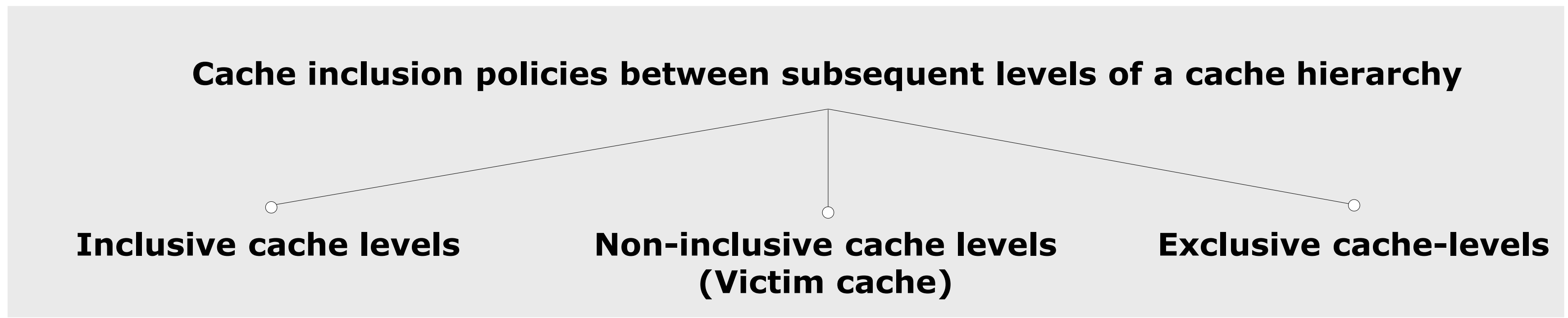
b) Re-architected L2/L3 cache hierarchy

b) Re-architected L2/L3 cache hierarchy [139]

- 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

*

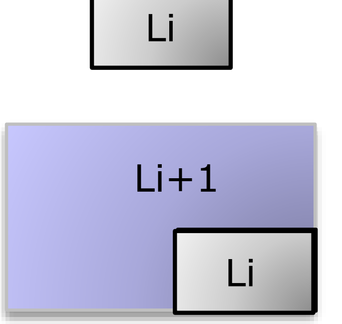
Changing the L3 cache inclusion policy from inclusive to non-inclusive -1

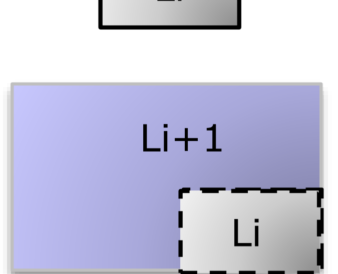


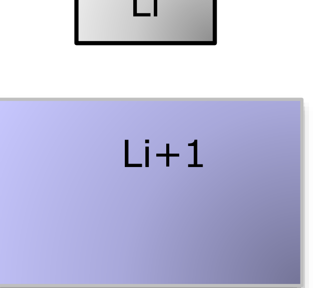
If the cache levels are inclusive, the higher cache level includes the content of the underneath cache level,
E.g. an inclusive L3 cache contains the content of the L2 cache

In non-inclusive cache levels the fact that a line is in level i does not imply that it is also in level i+1.

If the cache levels are exclusive, a data item (in fact a cache line) is contained either in the higher or in the lower level cache, but never in both.

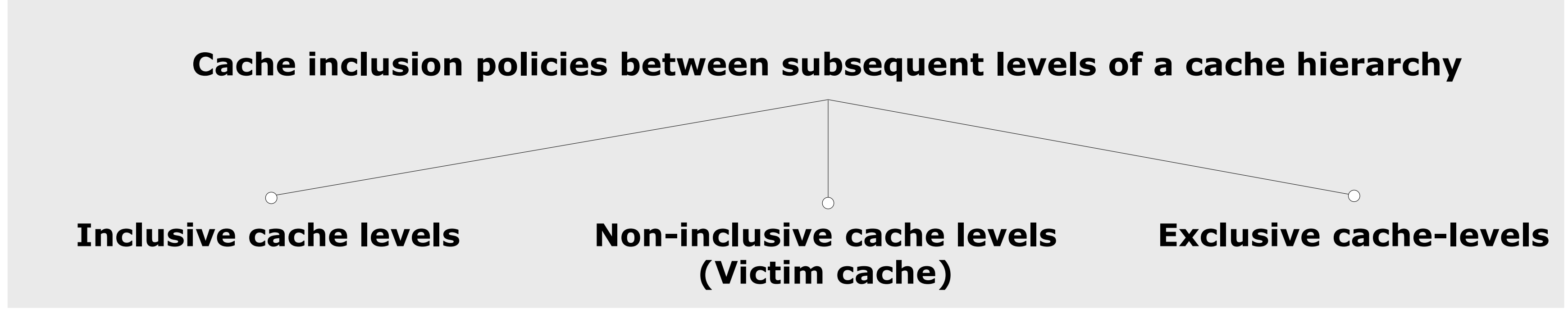






*

Assessing and use of cache inclusion policies



	Inclusive cache levels	Non-inclusive cache levels (Victim cache)	Exclusive cache-levels
Benefit	Inclusive cache levels reduce the snoop traffic, since in this case only the higher cache level needs to be snooped.	Non-inclusive cache levels seem to provide an efficient cache use.	Benefit of an exclusive cache levels is the efficient use of the cache space.
Drawback	Their drawback is the not efficient use of the cache space.	They need a more complex cache coherency protocol.	Their drawback is that they increase the snoop traffic since in this case both cache levels need to be snooped.
Examples	*Intel's L3 caches in their Haswell to Skylake lines*	*Intel's L2 caches in their Pentium 4 and Core 2 based lines* *Intel's L3 cache in their Skylake-SP line*	*AMD's L2 caches in their Athlon (K6) and Opteron based lines*

*

Inclusion policies of cache hierarchies in Intel's Core 2 lines [156]

	L2 to L1	L3 to L2	L4 to L3
Core 2/ Penryn	Shared Non-inclusive		
Nehalem/ Westmere	Private Non-inclusive	Shared Inclusive	
Sandy Bridge/ Ivy Bridge	Private Non-inclusive	Shared Inclusive Sliced	
Haswell Broadwell	Private Non-inclusive?	Shared Inclusive Sliced	Shared Non-inclusive
Skylake/ Kaby Lake	Private Non-inclusive?	Shared Inclusive Sliced	Shared Non-inclusive
Skylake-SP	Private Non-inclusive	Shared Non-inclusive Sliced	??

All caches are write-back caches (WB) (except of the L1 Instruction cache).
L3 cache tags show which L1 and/or L2 caches hold the cache line

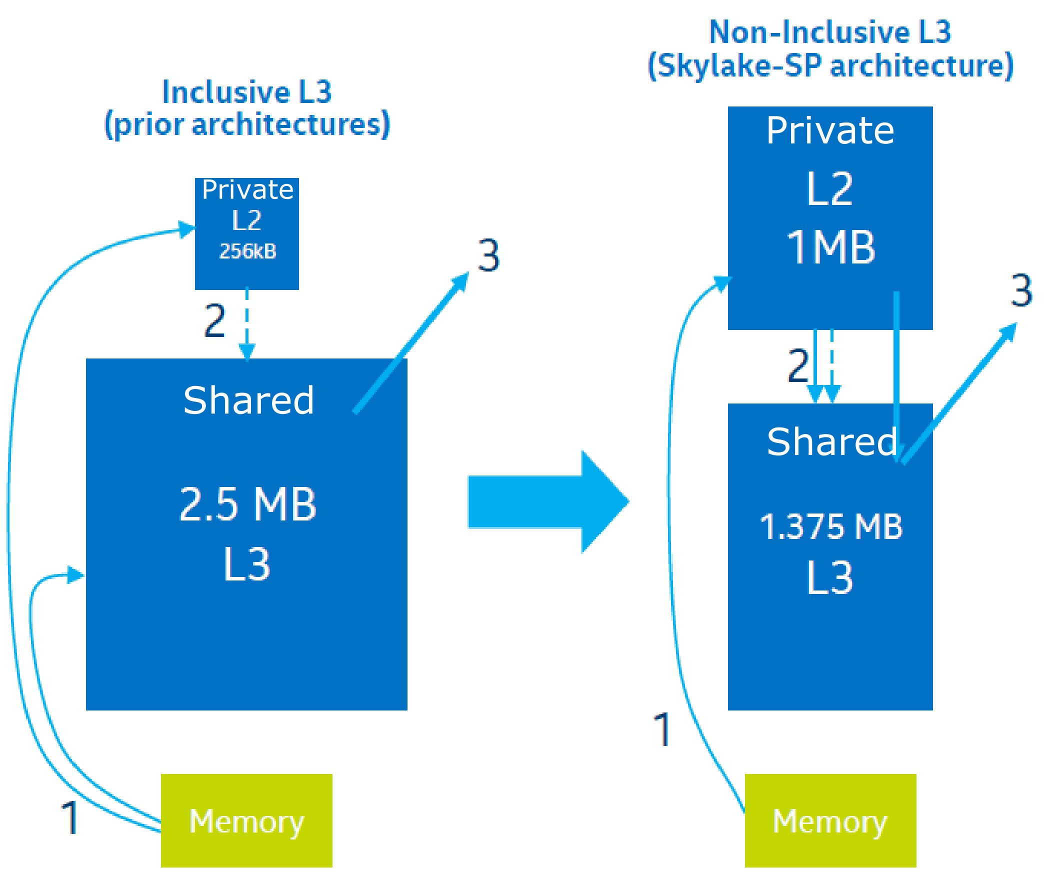
b) Re-architected L2/L3 cache hierarchy (5)

Changing the L3 cache inclusion policy from inclusive to non-inclusive -1

- In the previous Skylake generation the inclusive L3 cache amounts to up to 2.5 MB/core whereas the private L2 cache to 0.25 MB core and the L2 cache needs only about 10 % of the L3 cache space.
- By contrast, the Skylake-SP processor has an L3 cache of only 1.375 MB/core whereas the private L2 caches amount to 1 MB/core.
- Consequently, in the Skylake-SP processor the inclusive cache policy could not be used for the L3 cache, it had to be modified to the non-inclusive policy since then the L2 cache content is only partly included in the L3 cache.

*

Non-inclusive vs. inclusive L3 [139]

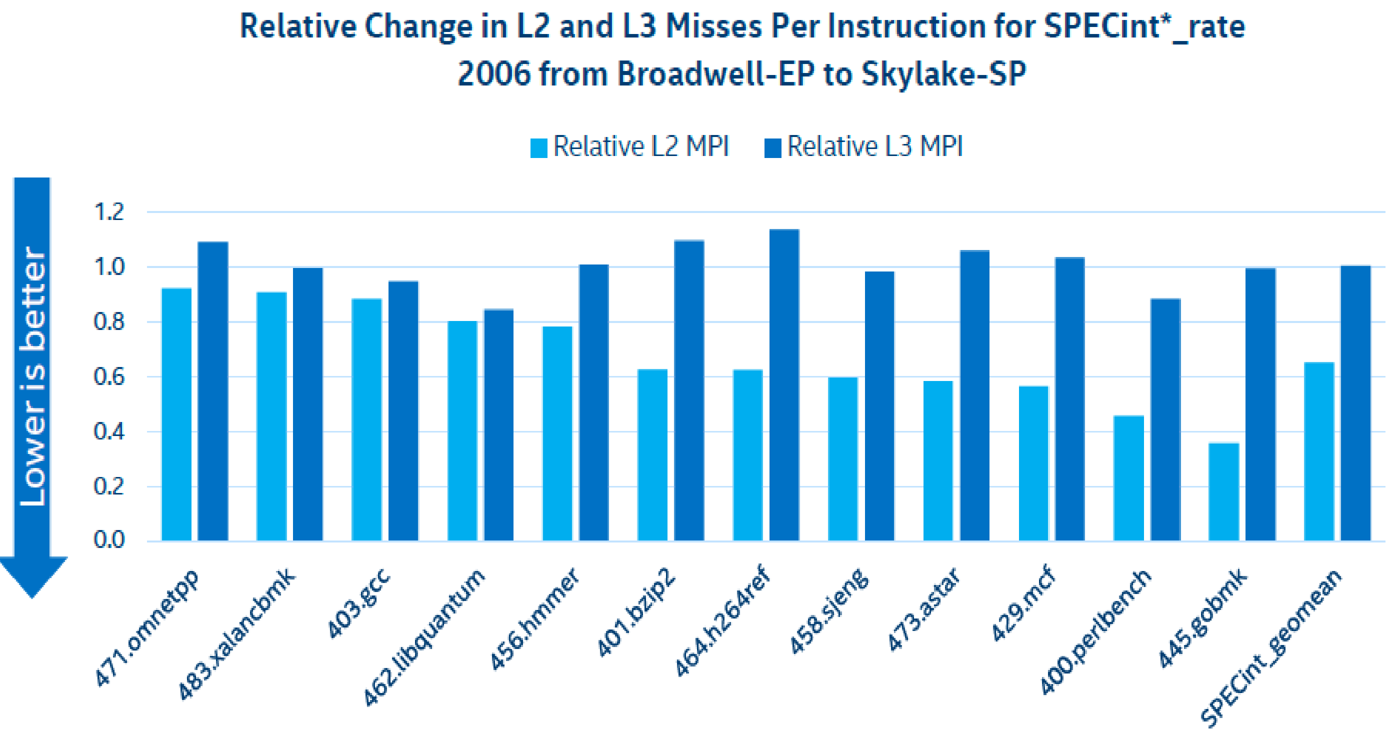


1. Memory reads fill directly to the L2, no longer to both the L2 and L3
2. When a L2 line needs to be removed, both modified and unmodified lines are written back
3. Data shared across cores are copied into the L3 for servicing future L2 misses

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

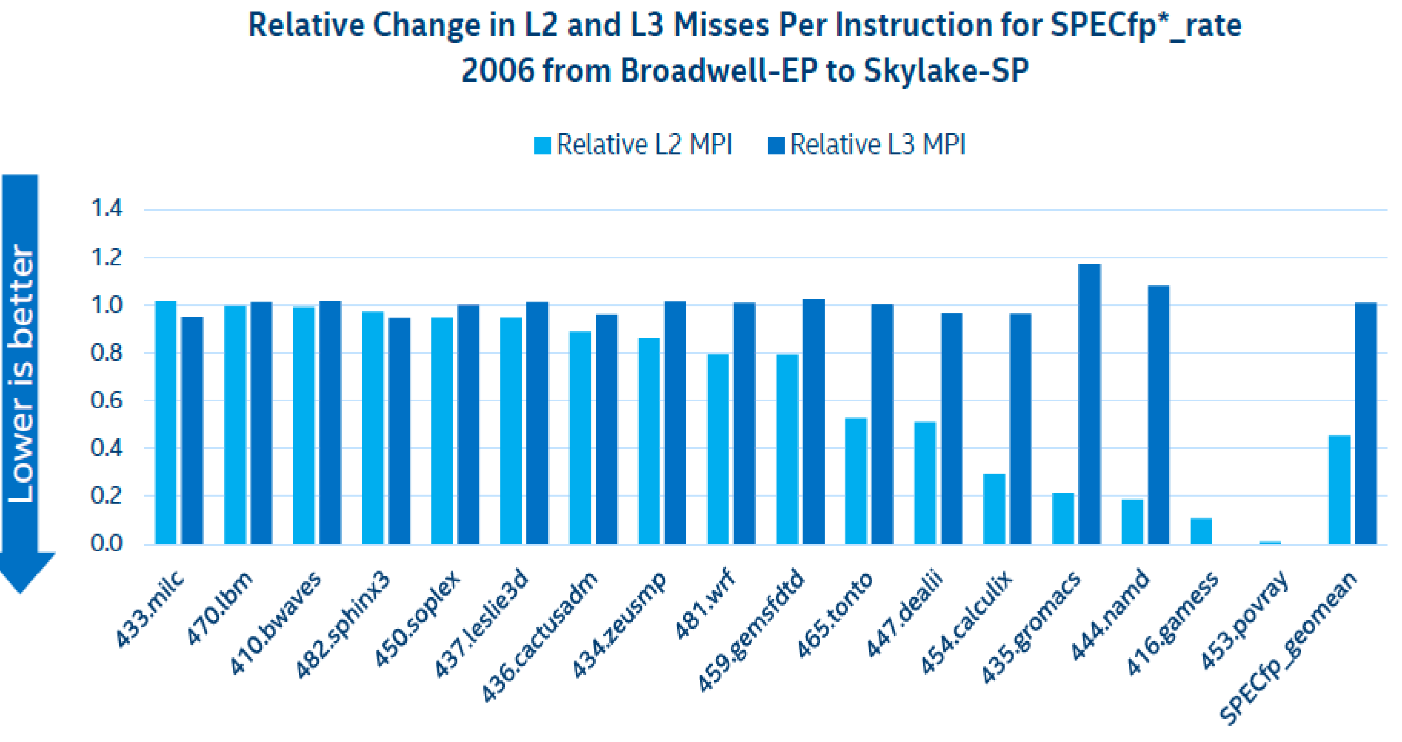
Cache misses in the Skylake-SP processor vs. the previous Broadwell-EP line -1 [139]



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

MPI: Misses Per Instruction

Cache misses in the Skylake-SP processor vs. the previous Broadwell-EP line-2 [139]

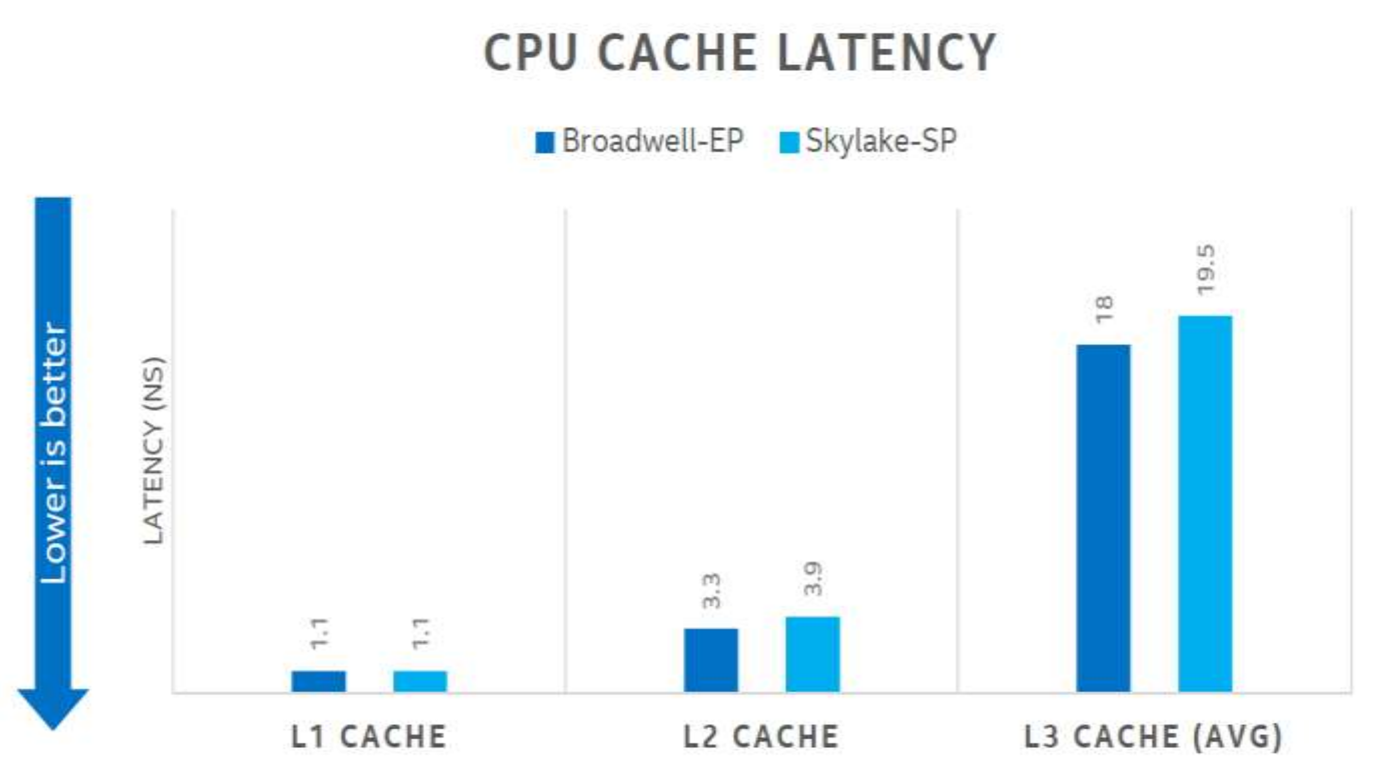


Skylake-SP cache hierarchy significantly reduces L2 misses without increasing L3 misses compared to Broadwell-EP

MPI: Misses Per Instruction

b) Re-architected L2/L3 cache hierarchy (9)

Skylake-SP's vs. Broadwell-EP cache latencies [139]



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

Example cache parameters of Intel's Skylake microarchitecture [156]

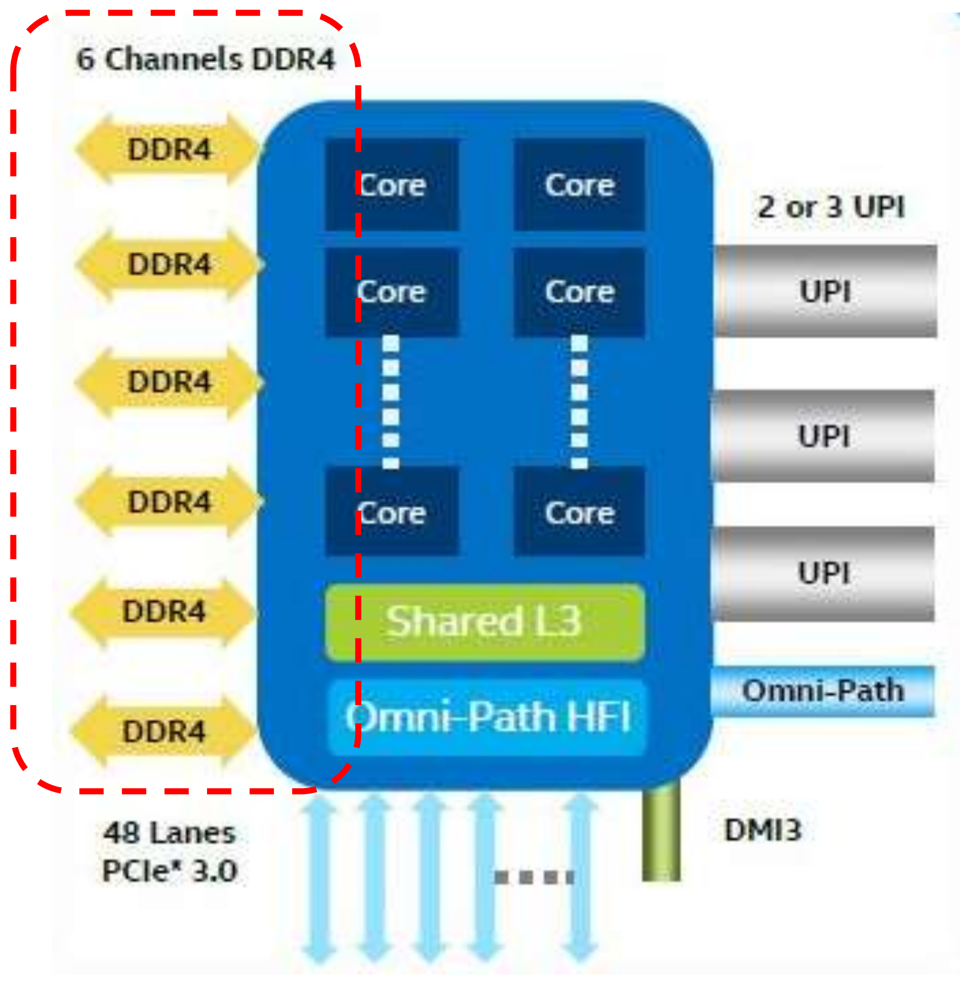
Cache Parameters of the Skylake Microarchitecture

Level	Capacity / Associativity	Line Size (bytes)	Fastest Latency[1]	Peak Bandwidth (bytes/cyc)	Sustained Bandwidth (bytes/cyc)	Update Policy
First Level Data	32 KB/ 8	64	4 cycle	96 (2x32B Load + 1*32B Store)	~81	Writeback
Instruction	32 KB/8	64	N/A	N/A	N/A	N/A
Second Level	256KB/4	64	12 cycle	64	~29	Writeback
Third Level (Shared L3)	Up to 2MB per core/Up to 16 ways	64	44	32	~18	Writeback

[1]Software-visible latency will vary depending on access pattern and other factors

c) 6 direct attached DDR4 memory channels

c) 6 direct attached DDR4 memory channels



Note that this kind of DRAM attachment is greatly different from that implemented in the previous Brickland 4S/8S platform.

*

Attaching memory to Intel's servers (from Nehalem to Broadwell based servers) -2

Attaching memory to Intel's servers

Direct attaching memory to EN/EP servers
(**EN:** Entry level, **EP**: Efficient performance)

Attaching up to 4 standard DDR memory channels direct to to processor die

Indirect attaching memory to EX servers
(**EX:** Extendable (high-end))

Attaching 8 standard DDR memory channels via 4 low line count serial or proprietary 64-bit parallel channels (SMI) with memory buffers, while two standard DDR memory channels can be connected to each memory buffer (SMB)

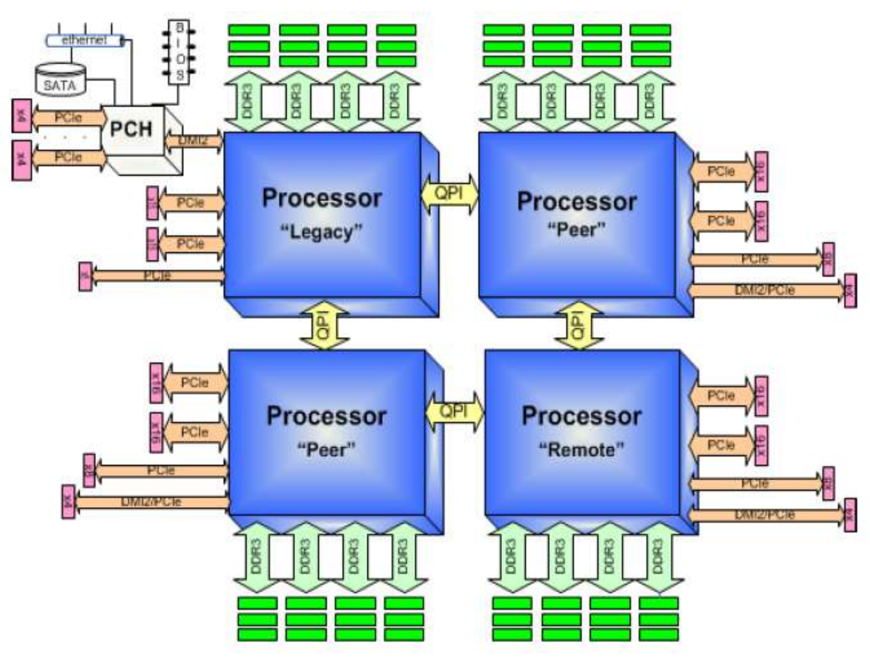


Figure: Sandy Bridge-EP based server [142]

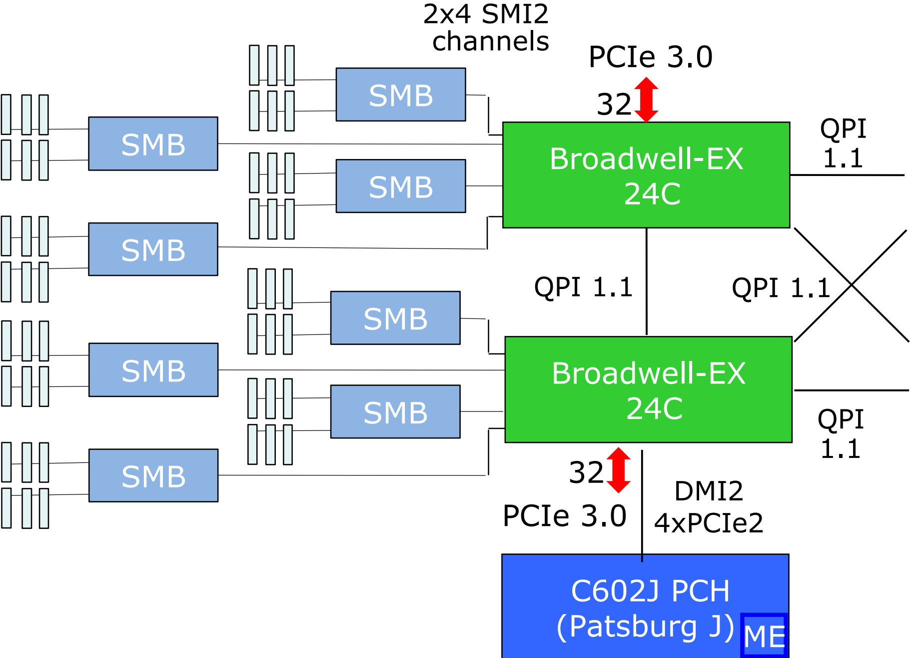


Figure: Part of a Broadwell-EX based server

*

c) 6 direct attached DDR4 memory channels (3)

Attaching memory to Intel's EP servers

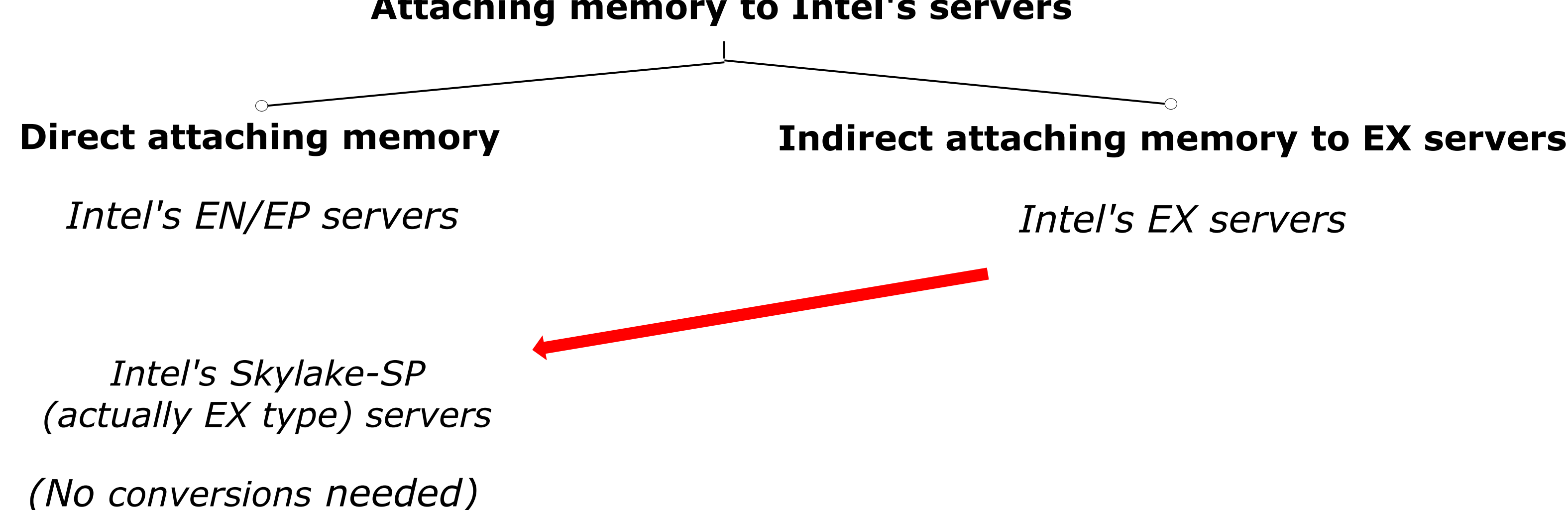
Line	Models	Core count	No. of memory channels	Kind of attachment	Socket	Intro
Nehalem-EP	E55xx	2C-4C	3xDDR3	Direct attached standard DDR3 or DDR4 memory channels	LGA 1366	2009
Westmere-EP	E5600	2C-6C	3xDDR3		LGA 1366	2010
Sandy Bridge-EP	E5-26xx E5-46xx	2C-8C 4C-8C	4xDDR3		LGA 2011	2012
Ivy Bridge-EP	E5-26xx v2 E5 46xx v2	4C-12C 4C-12C	4xDDR3		LGA 2011	2013 2014
Haswell-EP	E5-26xx v3 E5-46xx v3	4C-18C 6C-18C	4xDDR4		LGA 2011-3	2014 2015
Broadwell-EP	E5-26xx v4 E5-46xx v4	4C-22C 10C-22C	4xDDR4		LGA 2011-3	2016

c) 6 direct attached DDR4 memory channels (4)

Attaching memory to Intel's EX servers

Line	Models	Core count up to	No. of mem. channels/ socket	Kind of attachment	Socket	Intro
Nehalem-EX	E6500 (2S) E75xx 4S/8S)	8C	8xDDR3	4 low line count serial channels (SMI) with mem. buffers, 2 DDR3 channels/ mem. buffer	LGA 1567	2010
Westmere-EX	E7-2800 E7-4800 E7-8800	10C	8xDDR3		LGA 1567	2011
Ivy Bridge-EX	E7-28xx v2 E7-48xx v2 E7-88xx v2	15C	8xDDR3	4 low line count proprietary 64-bit parallel channels (SMI2) with memory buffers, 2 DDR3/4 channels/ mem. buffer	LGA 2011-1	2014
Haswell-EX	E7-48xx v3 E7-88xx v3	14C 18C	8xDDR3/4		LGA 2011-1	2015
Broadwell-EX	E7-48xx v4 E7-88xx v4	16C 22C	8xDDR3/4		LGA 2011-1	2016

Attaching memory to Intel's servers (from Nehalem to Broadwell based servers) -1



*

Sockets for direct attaching standard DDR memory channels to Intel's processors

P965 GMCH for Core 2 (2 cores)	**Westmere-EP** (up to 6 cores)	**Broadwell-EP** (up to 22 cores)	**Skylake-SP** (up to 28 cores)
34x34 mm	45x42.5 mm	58.5x51.0 mm	76x56 mm
BGA 1226	LGA 1366	LGA 2011-3	LGA 3647
2xDDR2 standard mem. channels on MCH	3xDDR3 standard mem. channels on processors	2x2 DDR4 standard mem. channels (2 on both side)	2x3 DDR4 standard mem. channels (3 on both side)

Note that physical/electrical constraints limit the number of attachable memory channels

Source of the pictures [144]

c) 6 direct attached DDR4 memory channels (7)

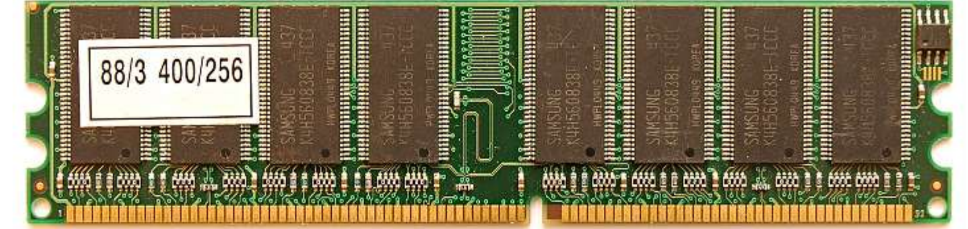
Pin counts of SDRAM to DDR4 DIMMs

All these DIMMs are 8-byte wide.

SDRAM (SDR)

168-pin

DDR

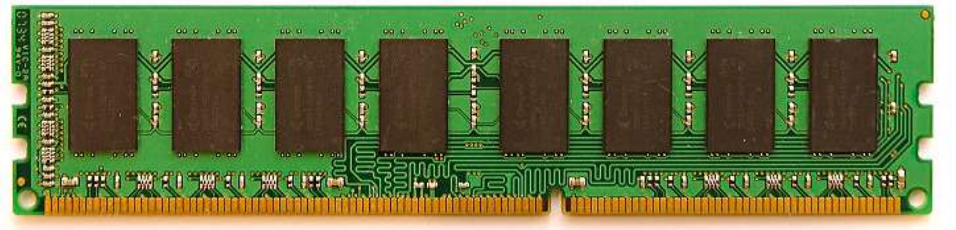


184-pin

DDR2

240- pin

DDR3



240-pin

DDR4

284-pin

*

Implementation of the memory subsystem -1 [139]

- 2 memory controllers with 3 channels each resulting in 6 DDR channels.
- Support of DDR4-2666, 2 DIMMs per channel

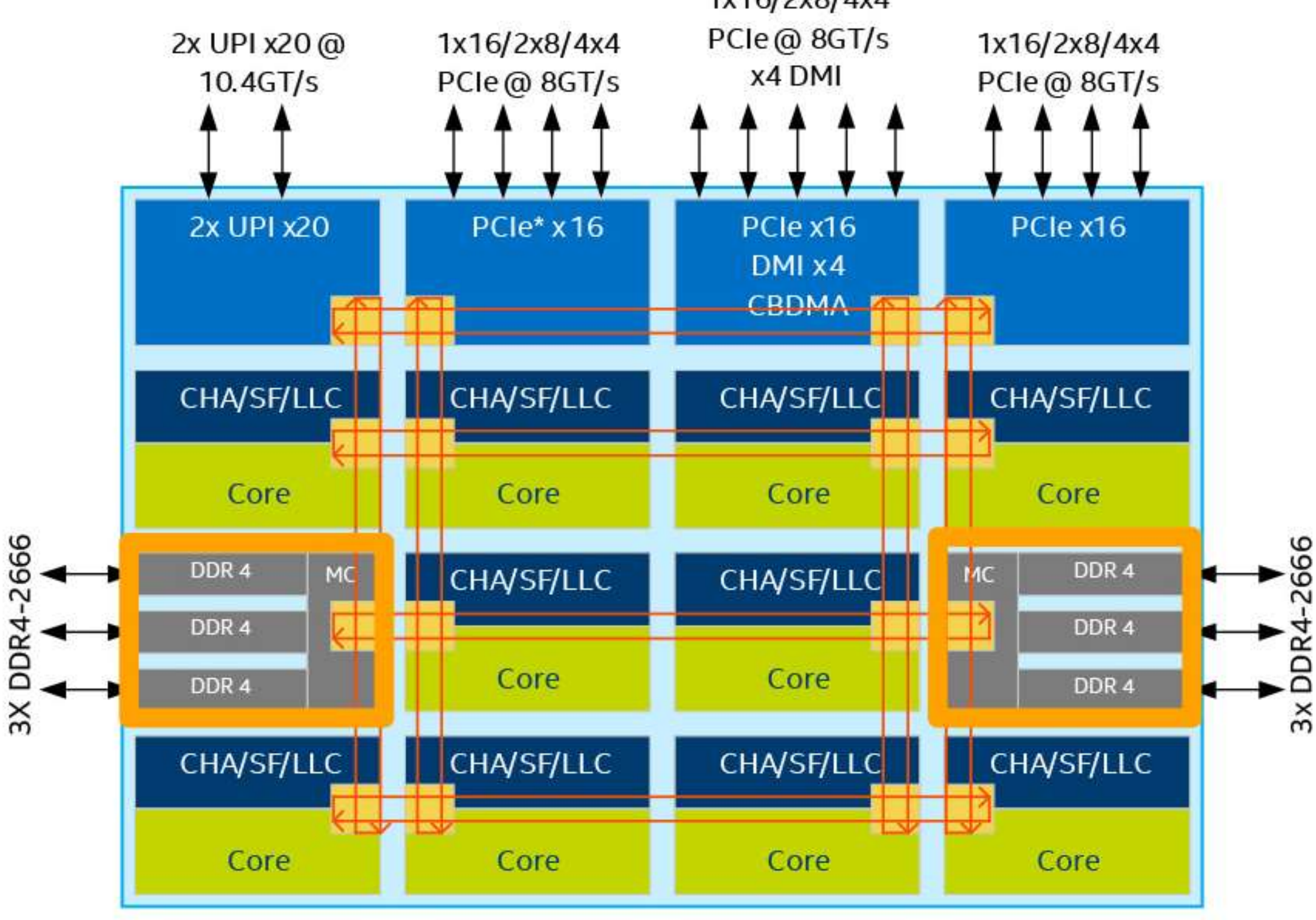


Figure: The memory subsystem [139]

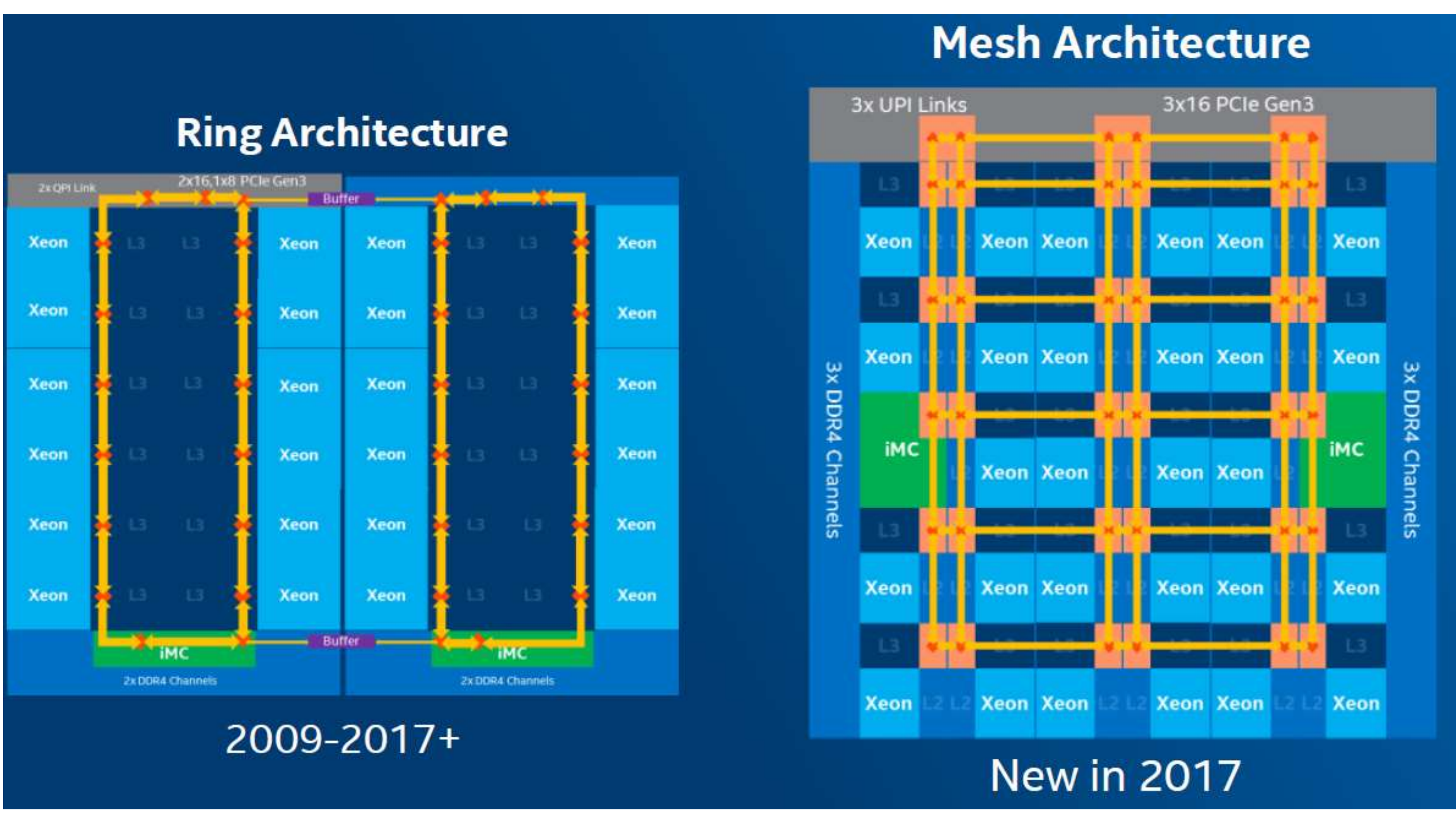
c) 6 direct attached DDR4 memory channels (9)

Implementation of the memory subsystem -2 [146]

- Up to 1.5 TB memory capacity per socket, assuming 2 128 GB DIMMs.
- > 60 % increase in memory bandwidth per socket compared to Xeon E5 v4 (Broadwell-EP)

d) Mesh architecture

d) Mesh architecture [143]

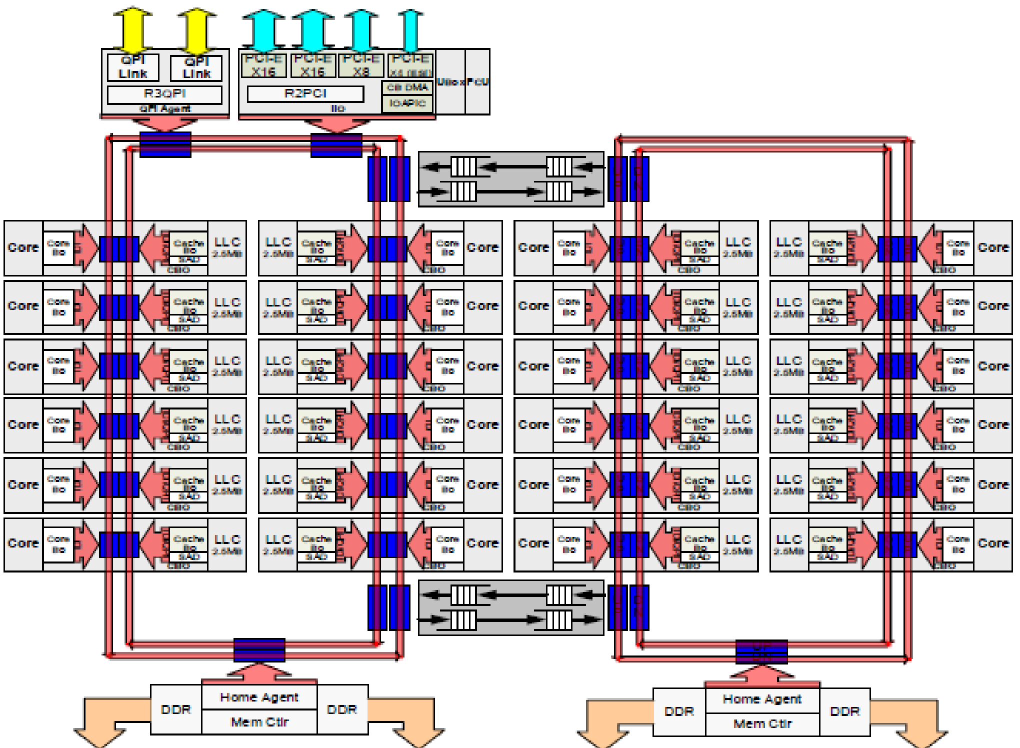


*

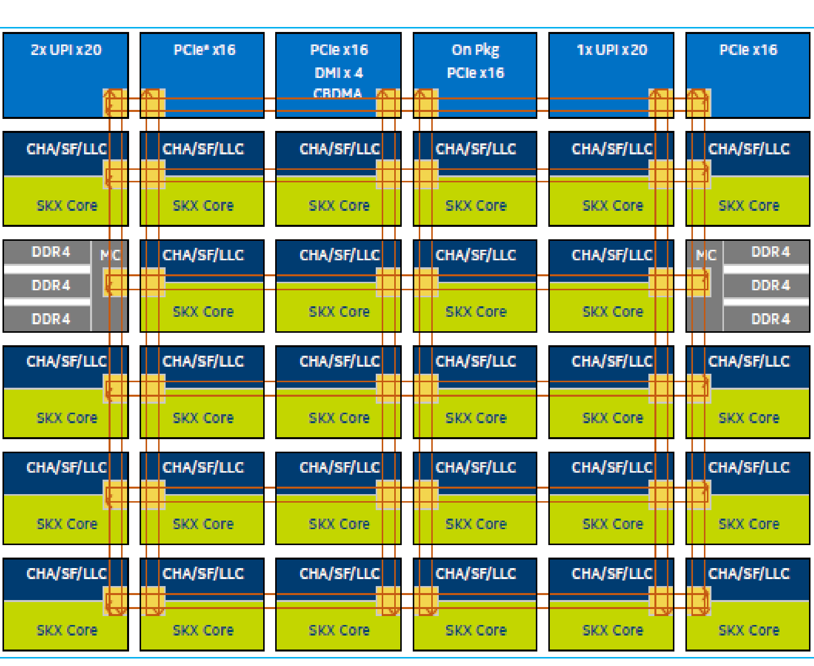
d) Mesh architecture (2)

Broadwell-EX's ring architecture vs. Skylake-SP's mesh architecture [139]

Broadwell EX 24-core die

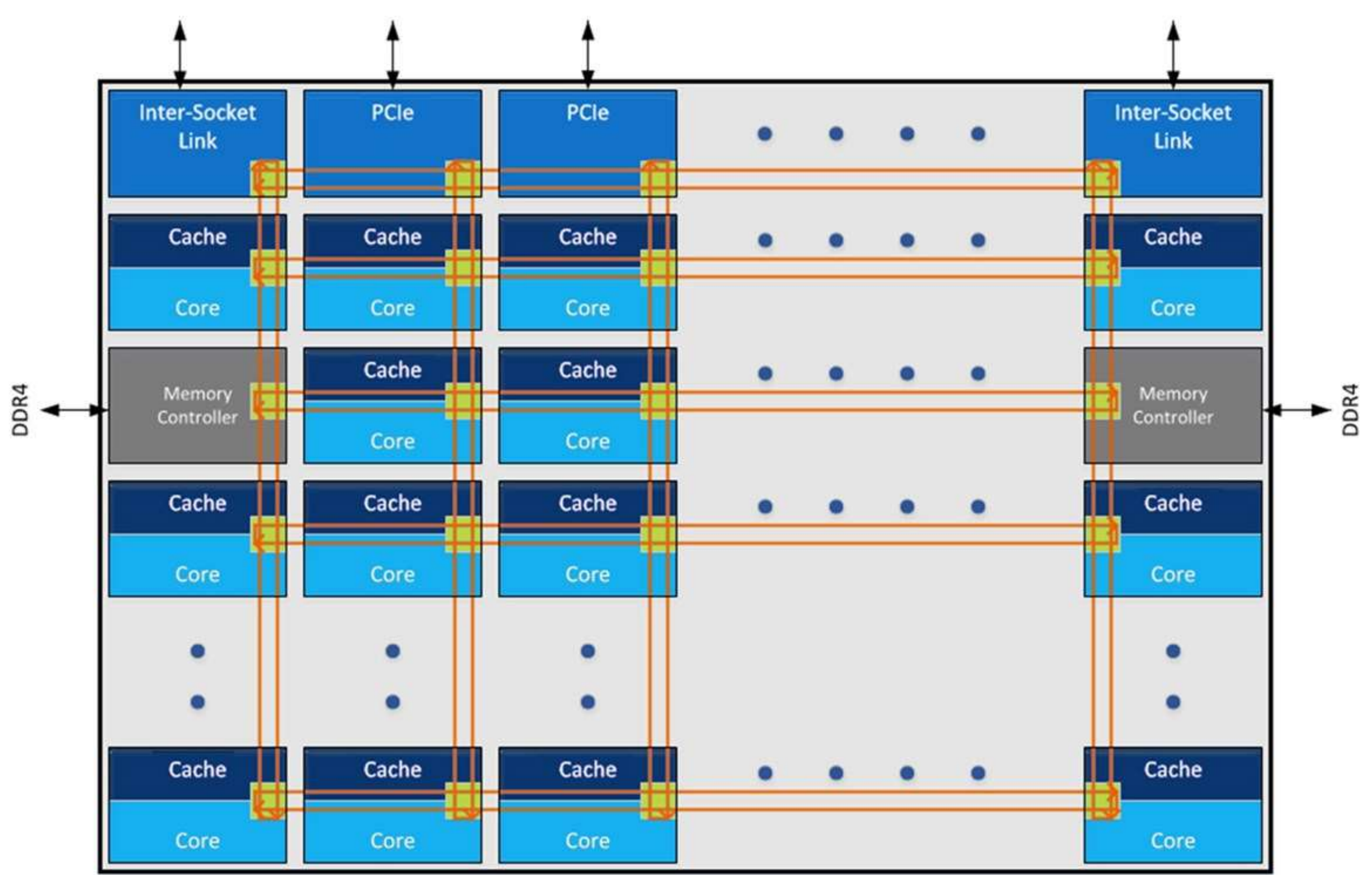


Skylake-SP 28-core die



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

Mesh architecture - more detailed [147]



d) Mesh architecture (4)

Benefits of the mesh architecture

- Higher performance
- Lower latencies
- Optimized for data sharing

*

Interconnection style of Intel's many core and multi-core processors

Interconnection style of Intel's many and multi-core processors

Ring architecture

Sony/Toshiba/IBM Cell (2006): 8 cores

Larrabee (2009). 24-32 cores (cancelled)
Xeon Phi
Knights Ferry (2010): 32 cores
Knights Corner (2012): 57-61 cores

Nehalem-EX (Beckton) (2010) up to 8 cores
All **Sandy Bridge** lines (2011) up to 8 cores
All **Ivy Bridge** lines (from 2012 on) up to 15 cores
to
All **Broadwell** lines (from 2014 on) up to 24 cores
Skylake (2015) up to 4 cores

Mesh (2D grid) architecture

Tile processor (2007): 80 cores
SCC (2010): 48 cores

Xeon Phi
Knights Landing (2016): up to 72 cores

Skylake-X (HED) (2017) up to 18 cores
Skylake-SP (2017) up to 28 cores

Note

While using a ring architecture in their 57 - 61 core Knights Corner processors presumably, Intel experienced very long and very different latencies for data sharing, this could led the firm to switch to a mesh architecture in their next Xeon Phi line.

Obviously, a mesh architecture has less different latencies when sharing data between cores than a ring architecture (see below).

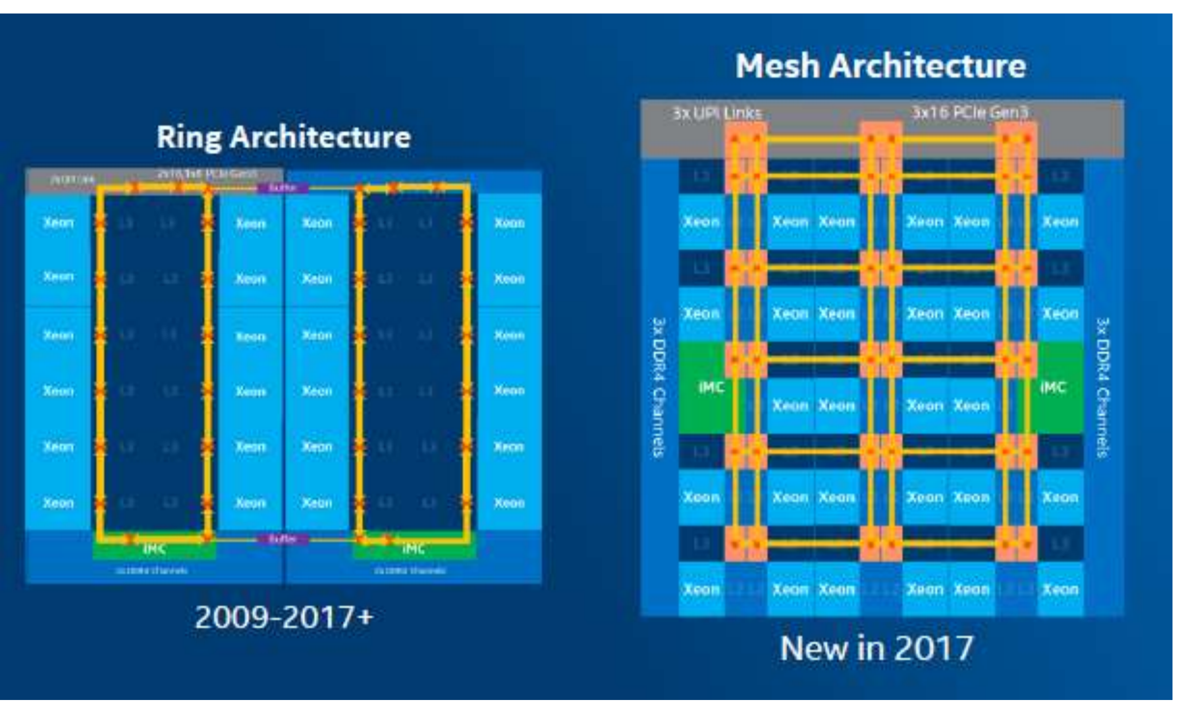


Figure: Contrasting ring and mesh architectures for data sharing between cores [143]

*

The layout of the ring interconnect on the Knights Corner die [157]

e) UPI (Ultra Path interconnect)

e) UPI (Ultra Path interconnect) [139]

- UPI replaces QPI.
- It provides faster links and improved message efficiency for the data packets, as seen below.

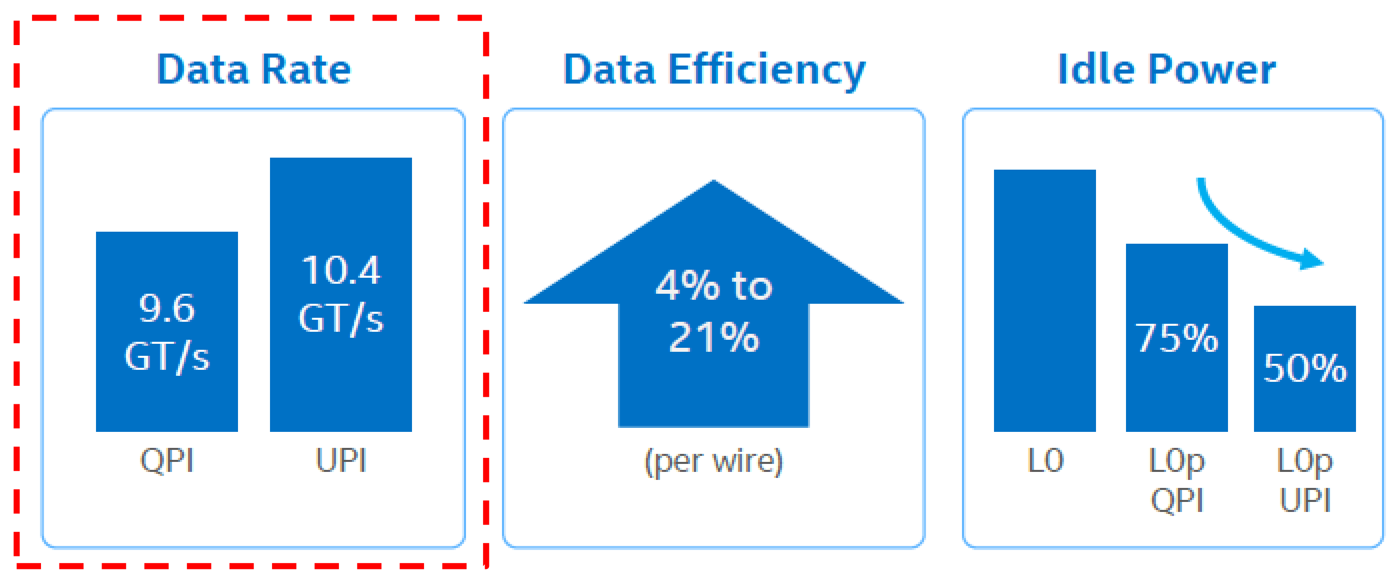


Figure: Improved data rate of the UPI interconnect vs. QPI {}

*

f) In-package integrated OmniPath host fabric interface

Will not be discussed

f) In-package integrated OmniPath host fabric interface -1 [146]

Implemented in the F-series of the Skylake-SP line, called also as the Skylake-F series.

The F-series are Platinum and Gold level (processors of the Skylake-SP line, as seen below.

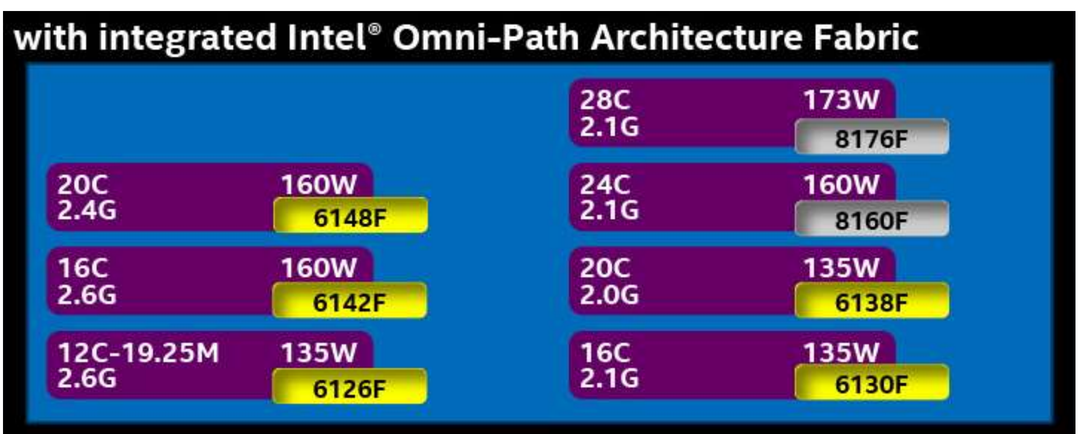


Figure: Skylake-SP models with integrated OmniPath host fabric interface [146]

*

f) In-package integrated OmniPath host fabric interface (2)

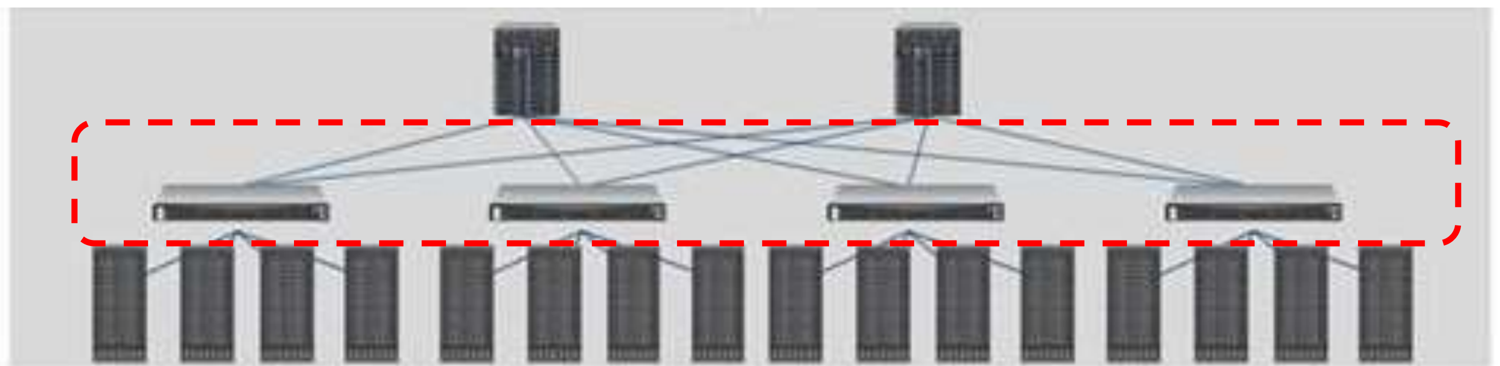
In-package integrated OmniPath host fabric interface -2 [146]

It aims at simplifying the implementation of processor clusters.

*

f) In-package integrated OmniPath host fabric interface (3)

Basic components of a processor cluster [158]



Storage system

Interconnect fabric (Switches)

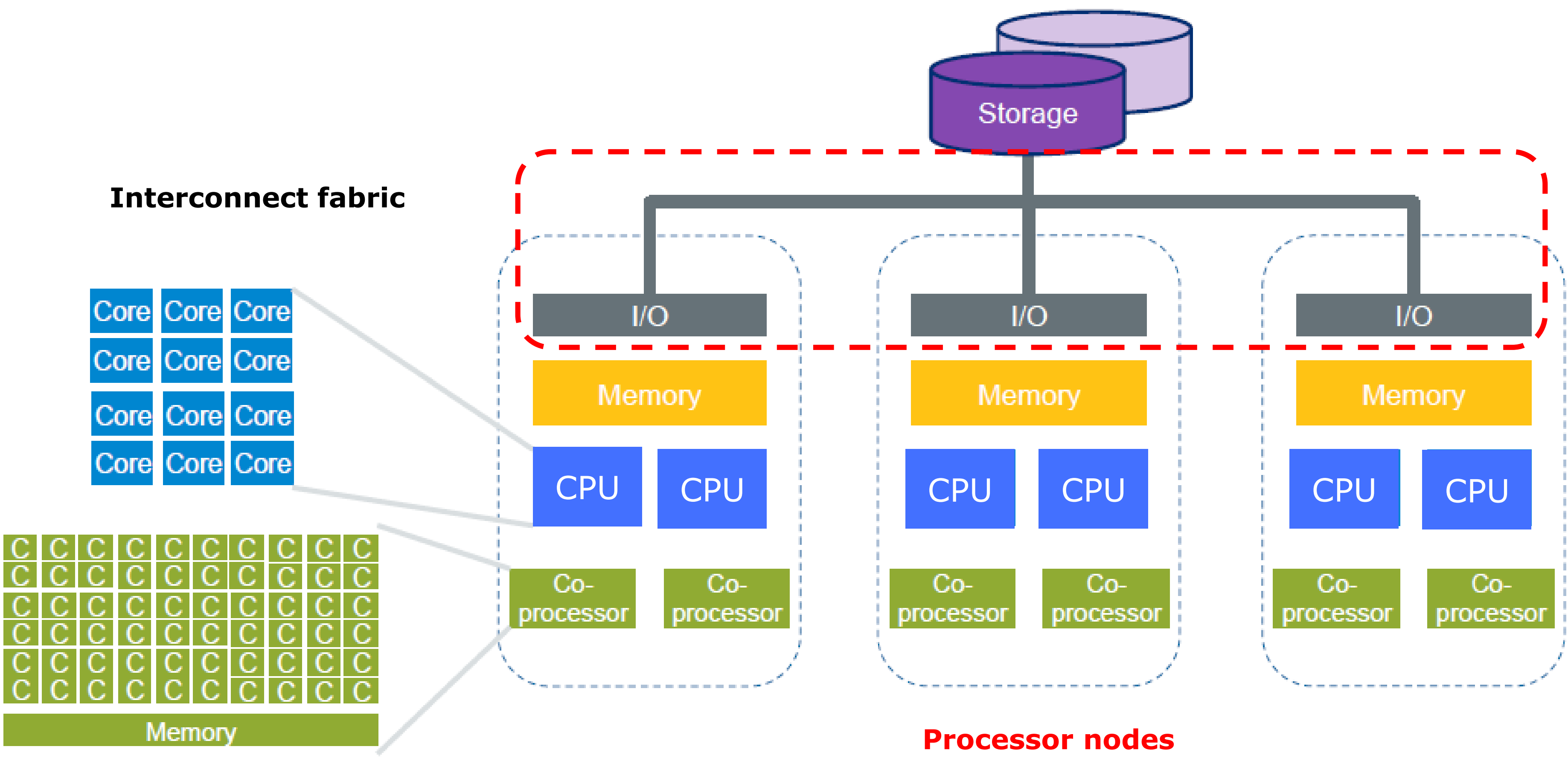
Processor nodes

Figure: Example of a processor cluster [158]

- Interconnect fabrics are used e.g. in processor clusters or supercomputers
- They interconnect processor nodes and also processor nodes with the storage system.

*

Possible layout of processor nodes [159]



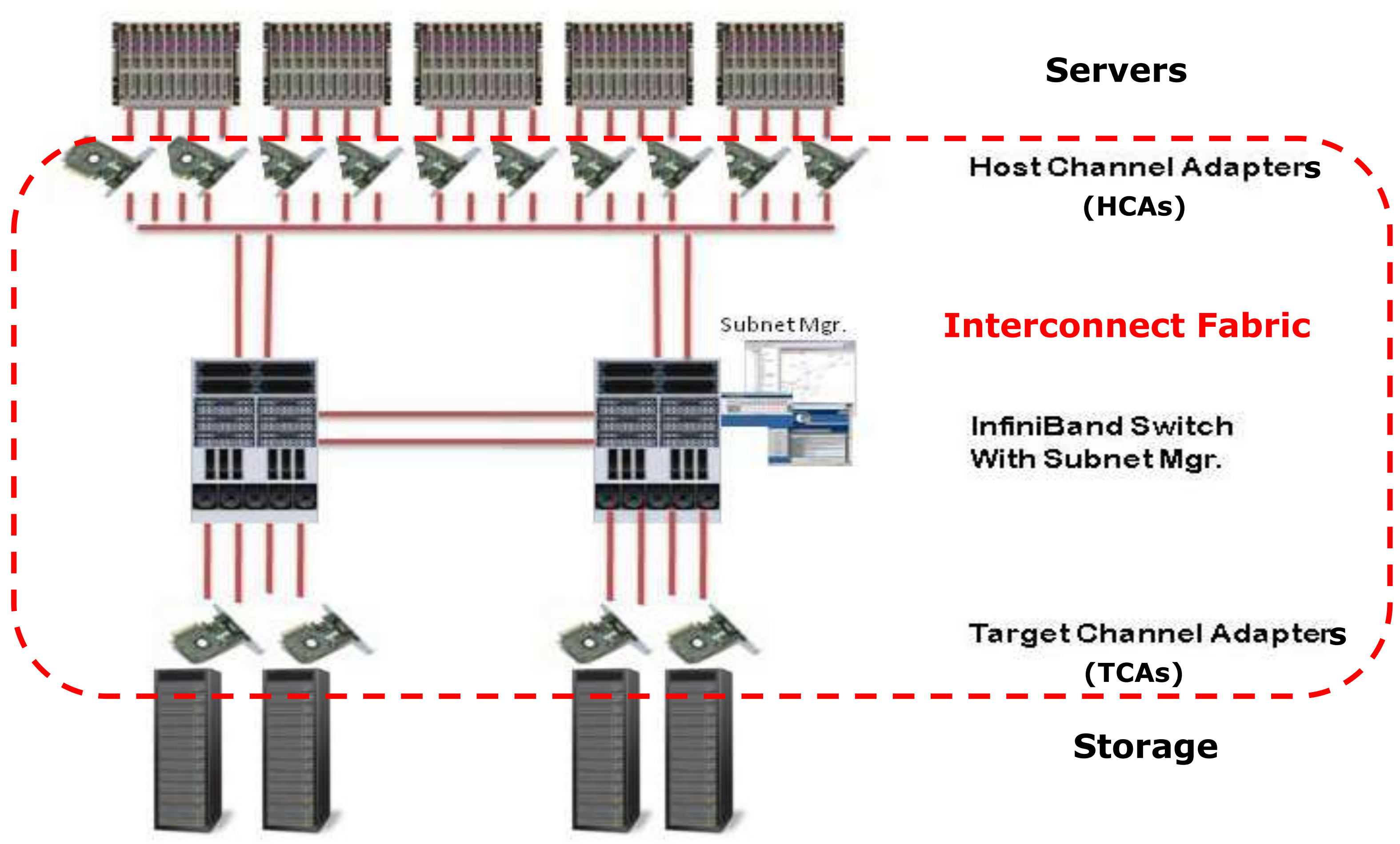
*

f) In-package integrated OmniPath host fabric interface (5)

Possible high speed interconnection technologies

- Ethernet based
- Fibre channels (FC) based
- InfiniBand (IB) based
- OmniPath based

Example: Processor cluster with InfiniBand based interconnect fabric and host and target channel adapters, called also fabric interfaces [160]



f) In-package integrated OmniPath host fabric interface (7)

Evolution of InfiniBand based high speed interconnect technologies to OmniScale

Infiniband (InfiniBand Trade Association. 1999)

↓

TrueScale (Qlogic 2008)
(HPC Enhanced version of InfiniBand)

↓

1/2012 Intel acquires QLogic's TrueScale business

↓

At the 2014 International Supercomputing Conference Intel announces both

- Knights Landing and
- the OmniScale interconnect fabric

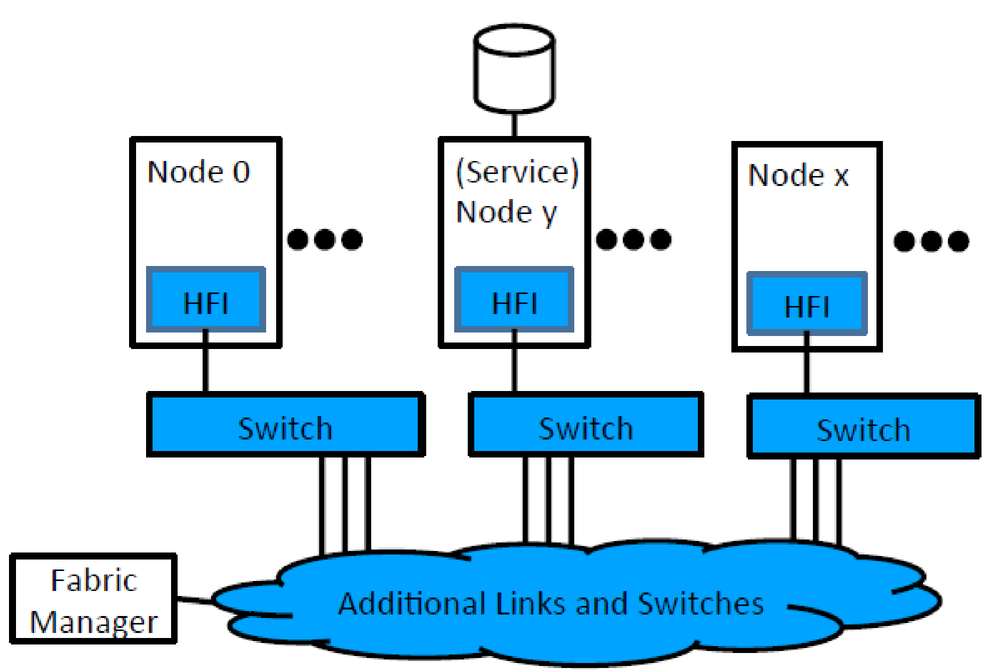
↓

11/2014 Intel renames OmniScale to OmniPath

↓

06/2016 Intel's Knights Landing line with in-package integrated OmniPath host channel adapter
07/2017 Intel's Xeon Skylake-SP line with in-package integrated OmniPath host channel adapter

Omni-Path Architecture Overview [167]



Omni-Path Components:

HFI – Host Fabric Interface

Provide fabric connectivity for compute, service and management nodes

Switches

Permit creation of various topologies to connect a scalable number of endpoints

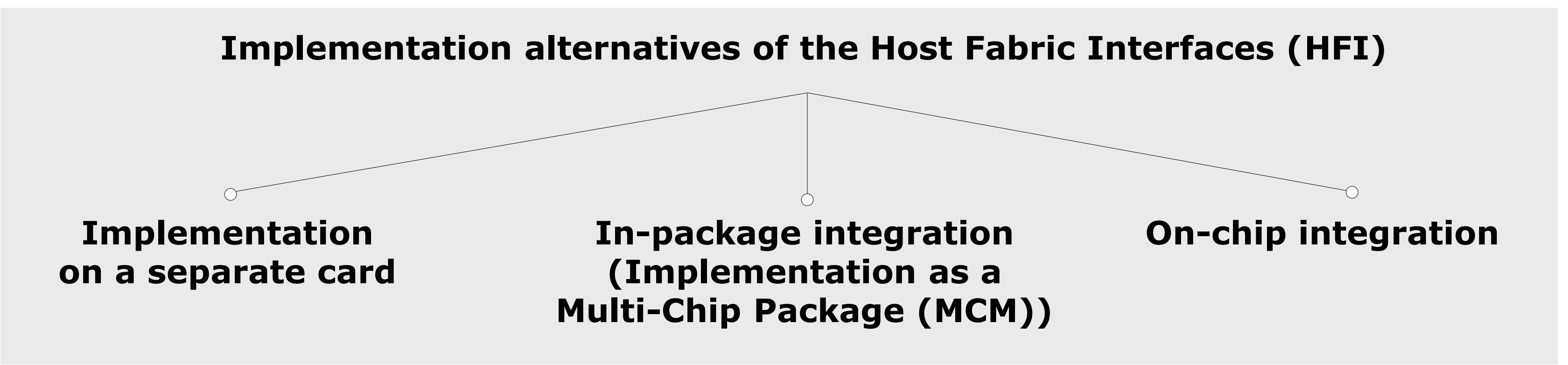
Fabric Manager

Provides centralized provisioning and monitoring of fabric resources

*

f) In-package integrated OmniPath host fabric interface (9)

Evolution of the implementation of Host Channel Adapters called Host Fabric Interfaces in OmniPath



Examples:

Previous systems with TrueScale

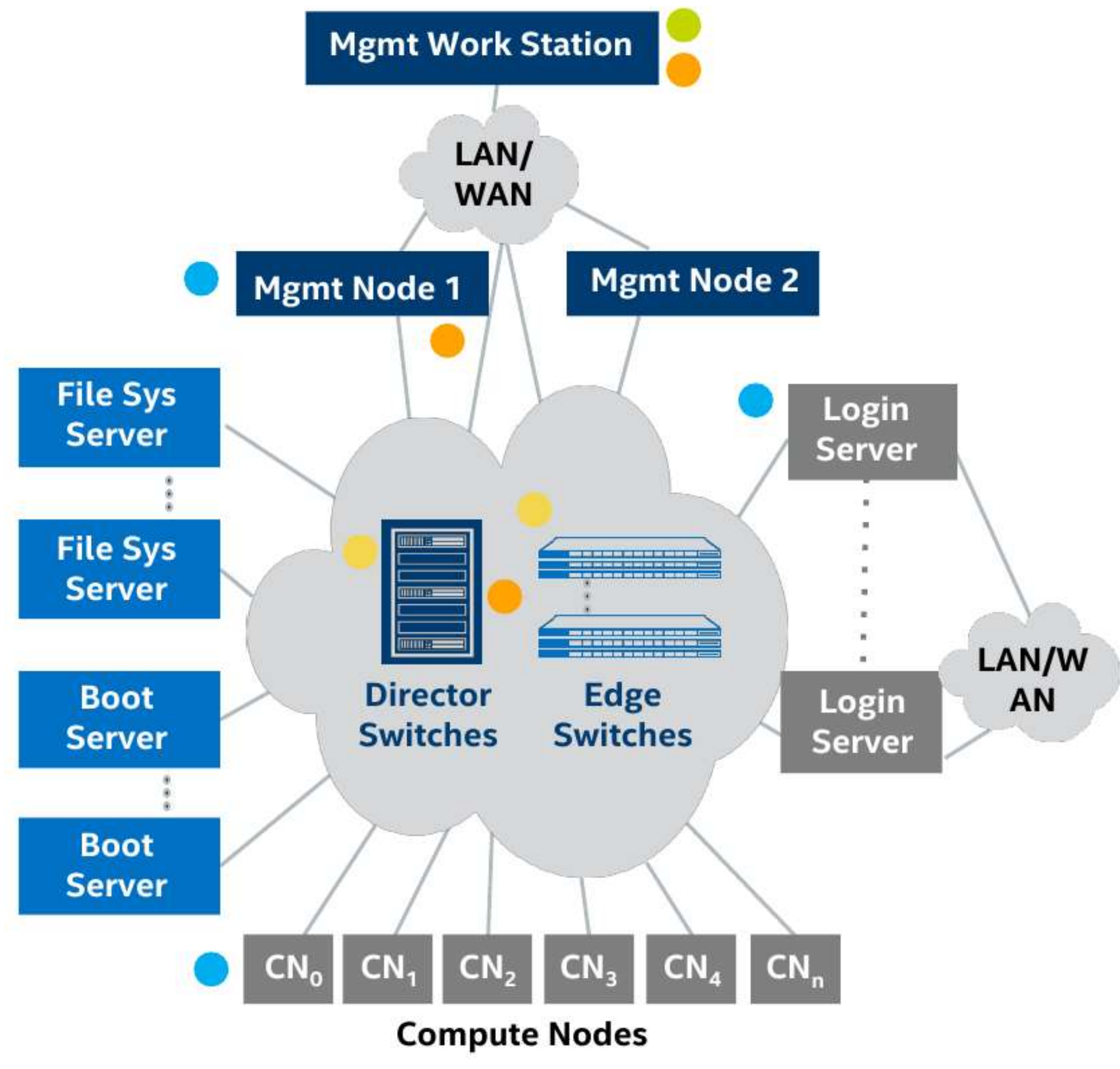
1. Gen. Knights Landing with OmniPath HCA

Skylake-F

2. Gen. Knights Landing with OmniPath HCA

*

System software needed for an interconnection fabric [161]



g) New socket (LGA 3647)

g) New socket (LGA 3647) [144]

Broadwell-EP
Broadwell-EX

LGA 2011

58.5x51.0 mm

Broadwell-EP

2x2 DDR4 standard
mem. channels
(2 on both side)

Broadwell-EX

4 SMI2 serial
memory links

Skylake-SP
Xeon Phy x200 (Knights Landing)

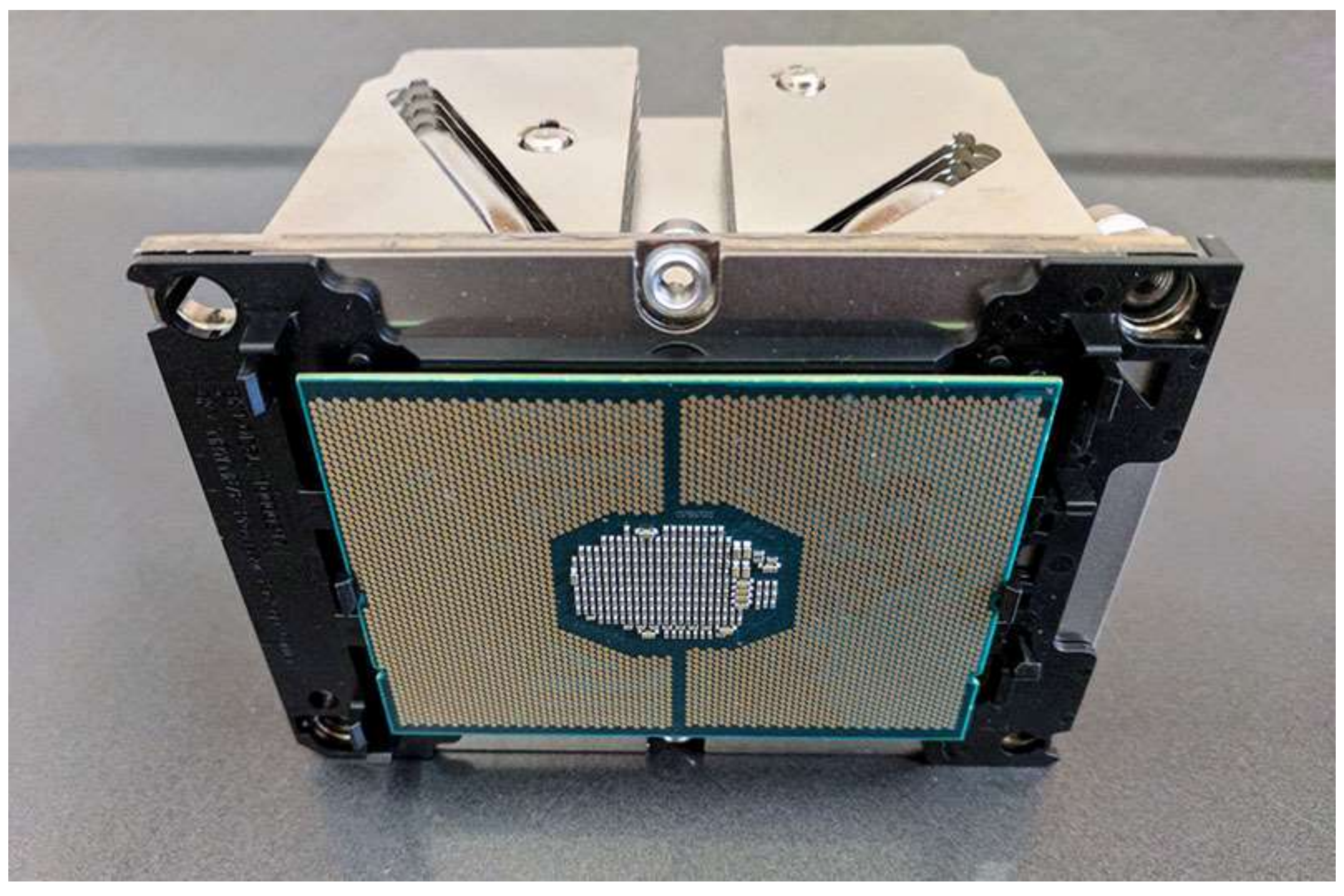
LGA 3647

76x56 mm

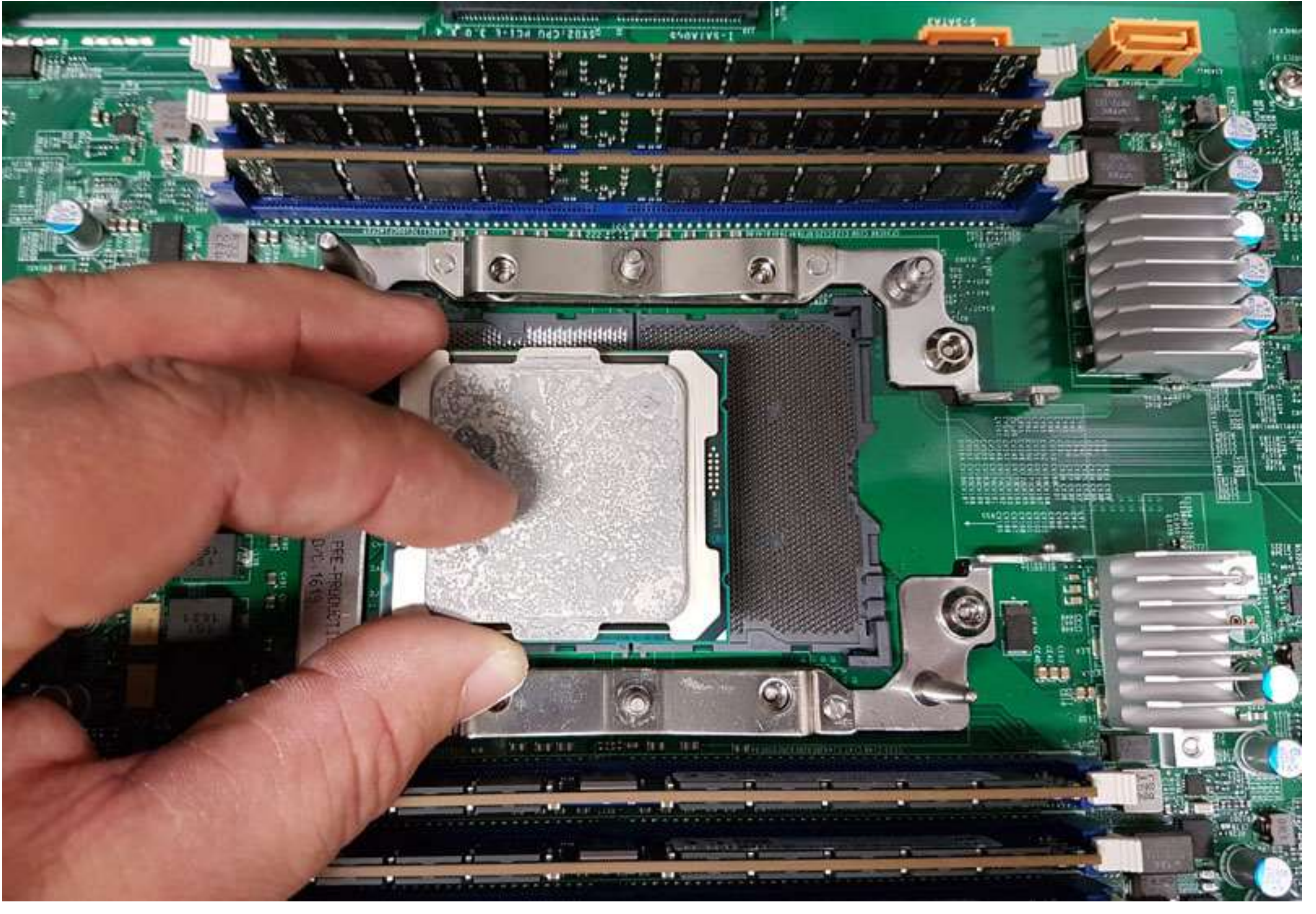
Skylake-SP

2x3 DDR4 standard
mem. channels
(3 on both side)

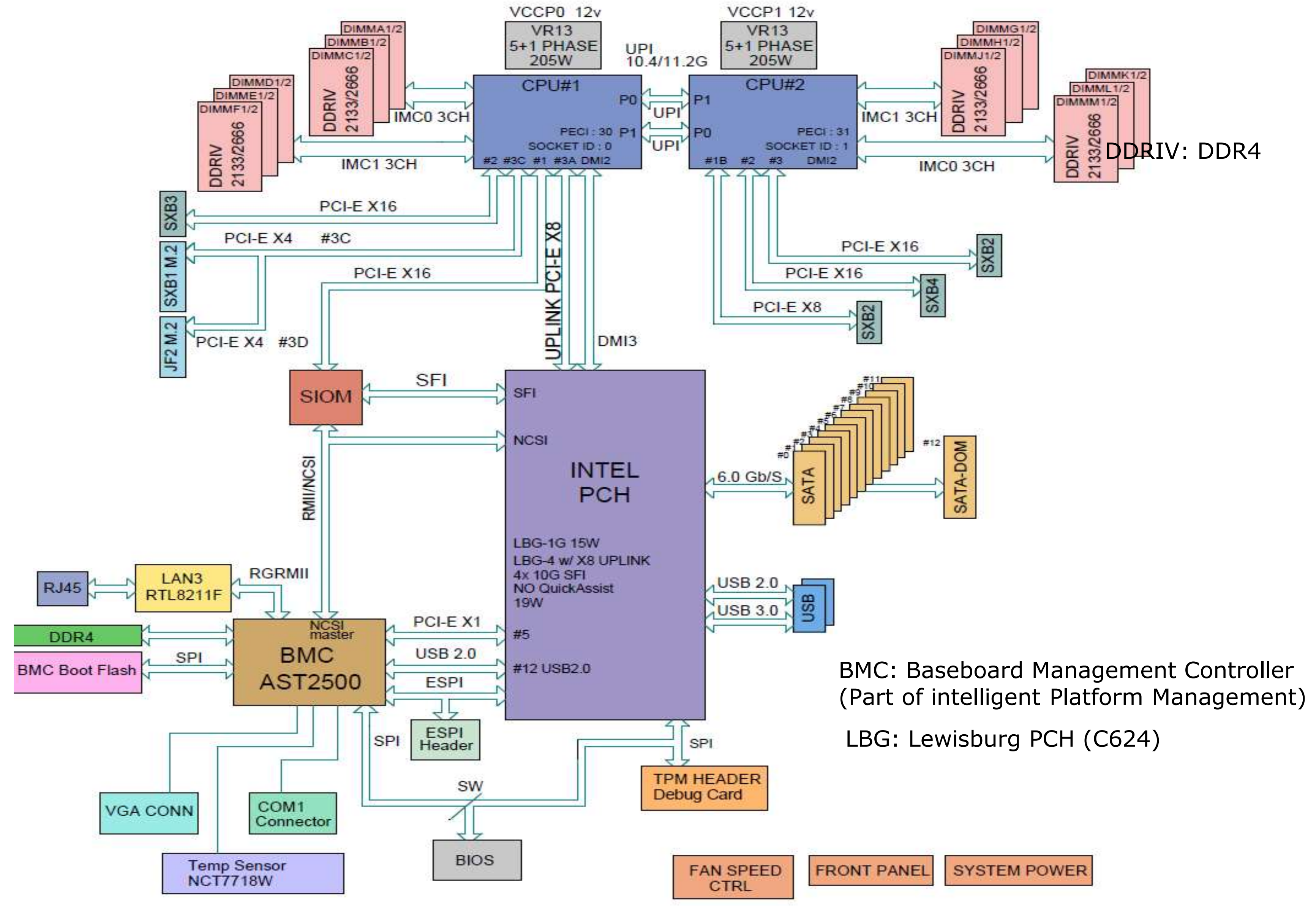
Skylake-SP socket with heatsink [144]



The Skylake-SP socket vs. the Broadwell-EP socket [144]



Block diagram of a Skylake-S based 2S server (Supermicro's X11DPT-B) [174]



DDRIV: DDR4

BMC: Baseboard Management Controller
(Part of intelligent Platform Management)

LBG: Lewisburg PCH (C624)

Motherboard of a Skylake-S based 2S server (Supermicro's X11DPT-B) [174]

4.3 The Cascade Lake processor line

- 4.3.1 Introduction
- 4.3.2 Key innovations of the Cascade Lake lines
- 4.3.3 The Cascade Lake-SP line
- 4.3.4 The Cascade Lake-AP line

4.3.1 Introduction

Intel's Xeon roadmap published in 08/2018 [176] -1

4.3.1 Introduction (2)

Intel's Xeon roadmap published in 08/2018 [176] -2

Note

- Although the Roadmap doesn't give a hint whether or not the Cascade Lake line belongs to the Purley platform, Intel disclosed at Hot Chips 2018 that Cascade Lake is compatible with the Purley platform [177].
- As key improvements of the 14 nm Cascade Lake line the roadmap points out:
  - the VNNI ISA extension for supporting DL (Deep Learning)
  - the Optane persistent memory and
  - security mitigations.

4.3.2 Key innovations of the Cascade Lake lines

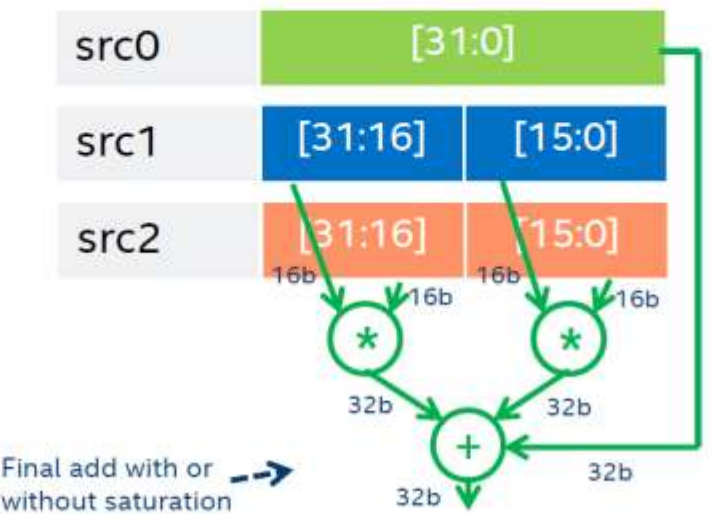
4.3.2 Key innovations of the Cascade Lake lines

a) Introduction of the VNNI ISA extension

VNNI (Vector Neural Network Instructions) provide 8 and 16-bit integer operations for Deep Learning applications.

Example: The VNNI-16 (Variable precision Vector Neural Network Instruction) [54]

- Vector Neural Network Instructions
- Variable precision
  - ***Inputs***: 16-bit INT
  - ***Outputs***: 32-bit INT
- Variable precision is best of both worlds
  - **Same operations/instruction as 'half precision'**
    - 2x OPS vs Single Precision
  - **Similar output precision for optimal training convergence**
    - 31 bits of INT32 vs 24 bits of mantissa in FP32
  - The obvious trade-off is the associated overhead on handling dynamic range in software (fixed precision)



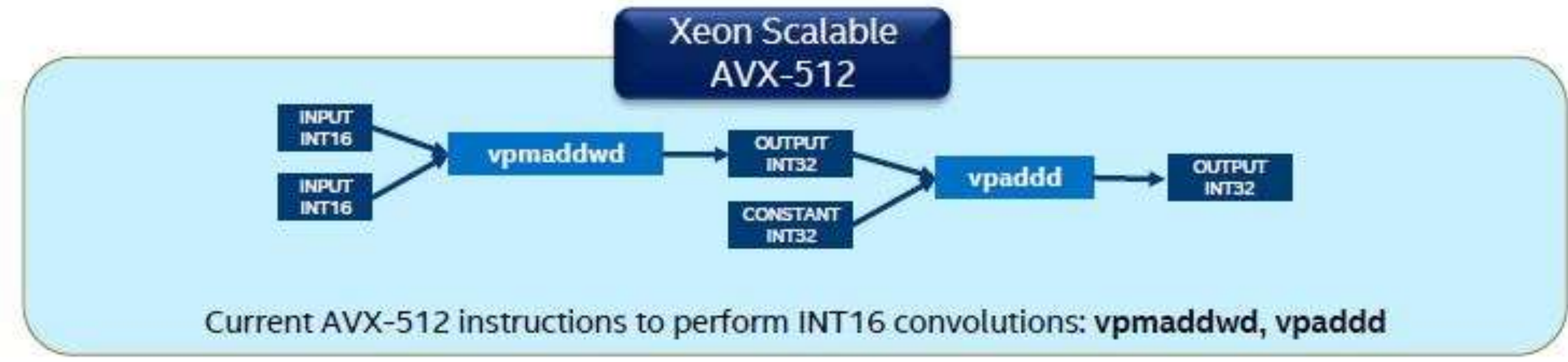
Operation

Dual 16-bit integer operands are multiplied, 32-bit results are added along with the scr0 operand.

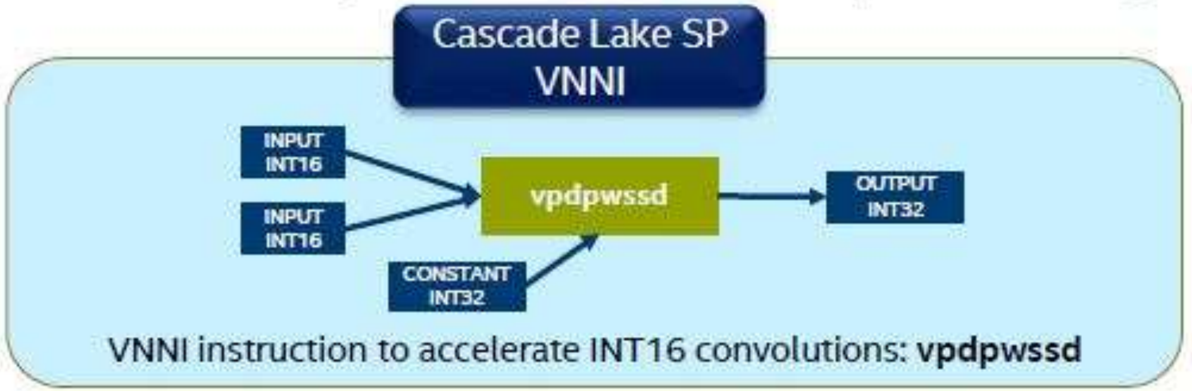
4.3.2 Key innovations of the Cascade Lake lines (3)

Speeding up the execution of the INT16xINT16 + INT32 operation by VNNI [177]

AI/DL Inference Enhancements on INT16 with VNNI

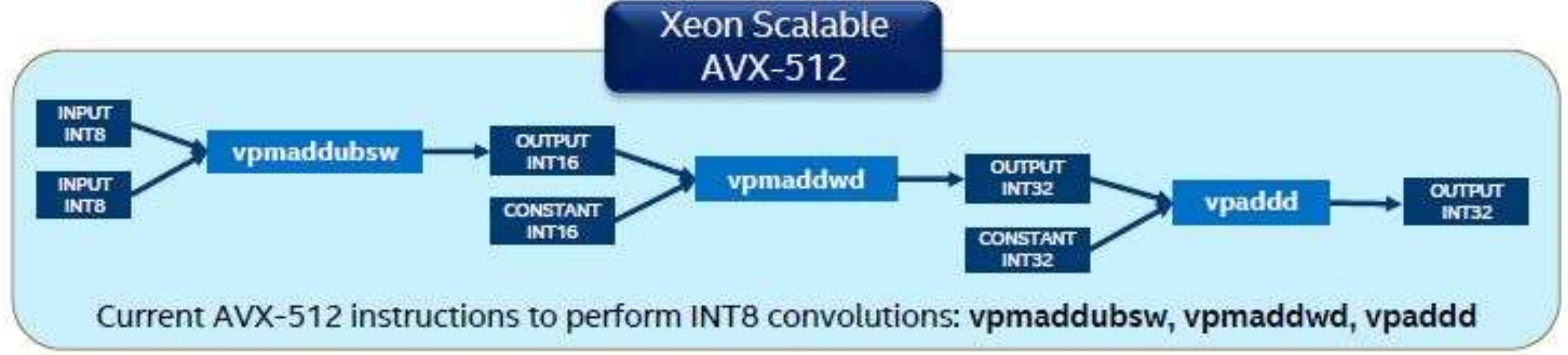


New instructions for accelerating AI on Intel® Xeon® Scalable processors using int16 data

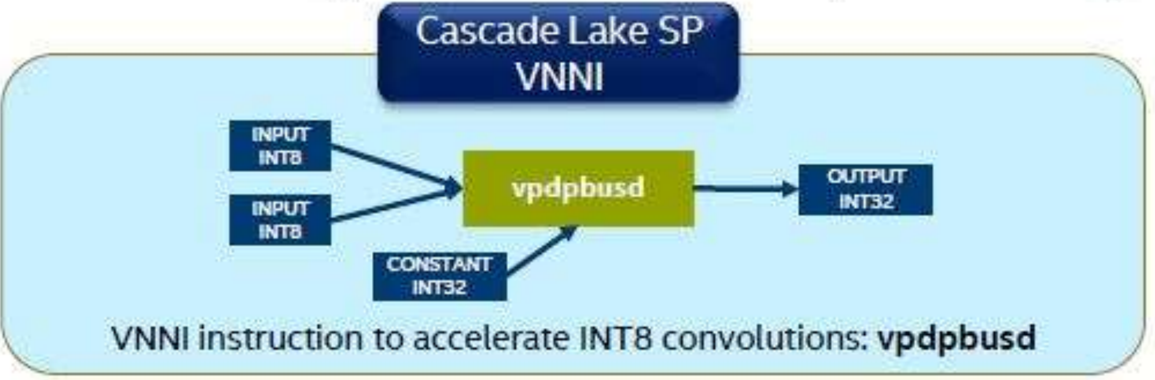


Speeding up the execution of the INT8xINT8 + INT32 operation by VNNI [177]

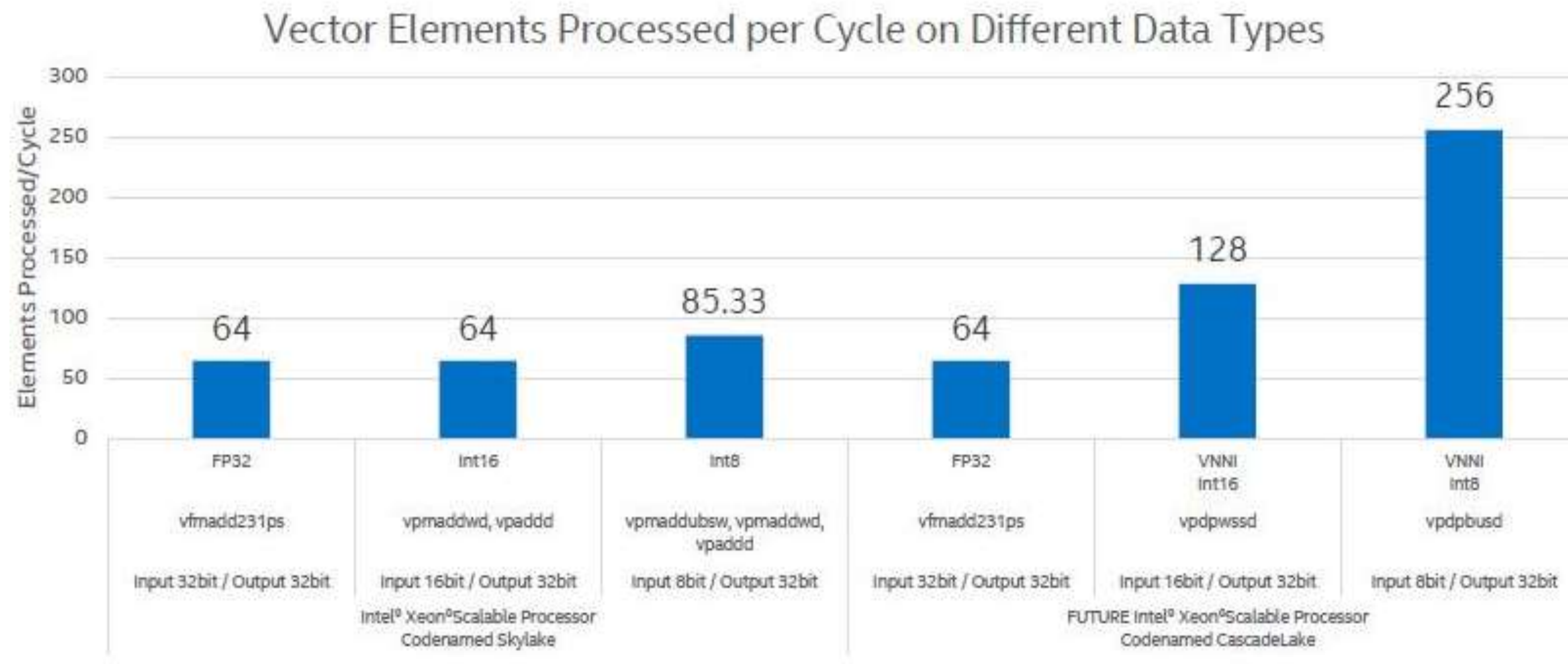
AI/DL Inference Enhancements on INT8 with VNNI



New instructions for accelerating AI on Intel® Xeon® Scalable processors using int8 data



Per core throughput per cycle while processing different data types with VNNI [177]



With speeding up INT8 and INT16 computations by VNNI Intel hopes to supersede low power GPU accelerators in inference computations [177].

b) Intel Optane DC Persistent Memory -1

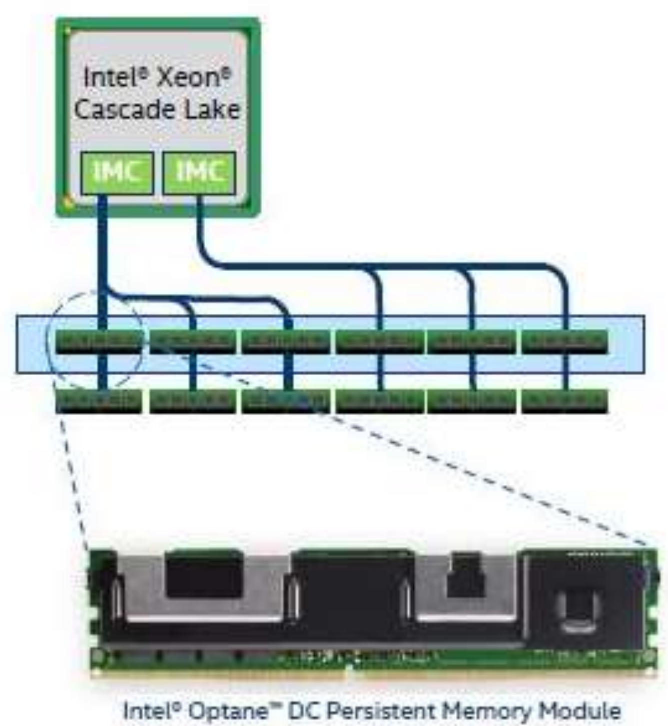
It is Optane memory mounted onto a DDR4 DIMM with DDR4 compatible contacts, as seen below.

Figure: Optane DC Persistent Memory [180]

4.3.2 Key innovations of the Cascade Lake lines (7)

b) Intel Optane DC Persistent Memory -2

Optane DC Persistent Memory is directly attached to the memory controller instead of the PCIe bus, as indicated below.



- DDR4 electrical and physical interface with proprietary protocol extensions
- Memory channel can be shared between DDR4 and Intel® Optane™ DC persistent memory modules
  - Enables systems to support greater than 3TB of system memory per CPU socket
- Cache line size accesses
- Idle latency close to DDR4 DIMMs

Figure: Connecting Intel's Optane DC Persistent Memory to a processor [177]]

Remarks to the Optane memory

- It is based on the 3D XPoint memory technology, announced by Intel and Micron in 2016.
- 3D Xpoint (Cross-Point) memory is a high-density, stackable, bit-level addressable matrix of non-volatile memory, as indicated below.
- It provides more endurability than SSD devices.
- First devices with Optane memory has been introduced in 1H 2017.

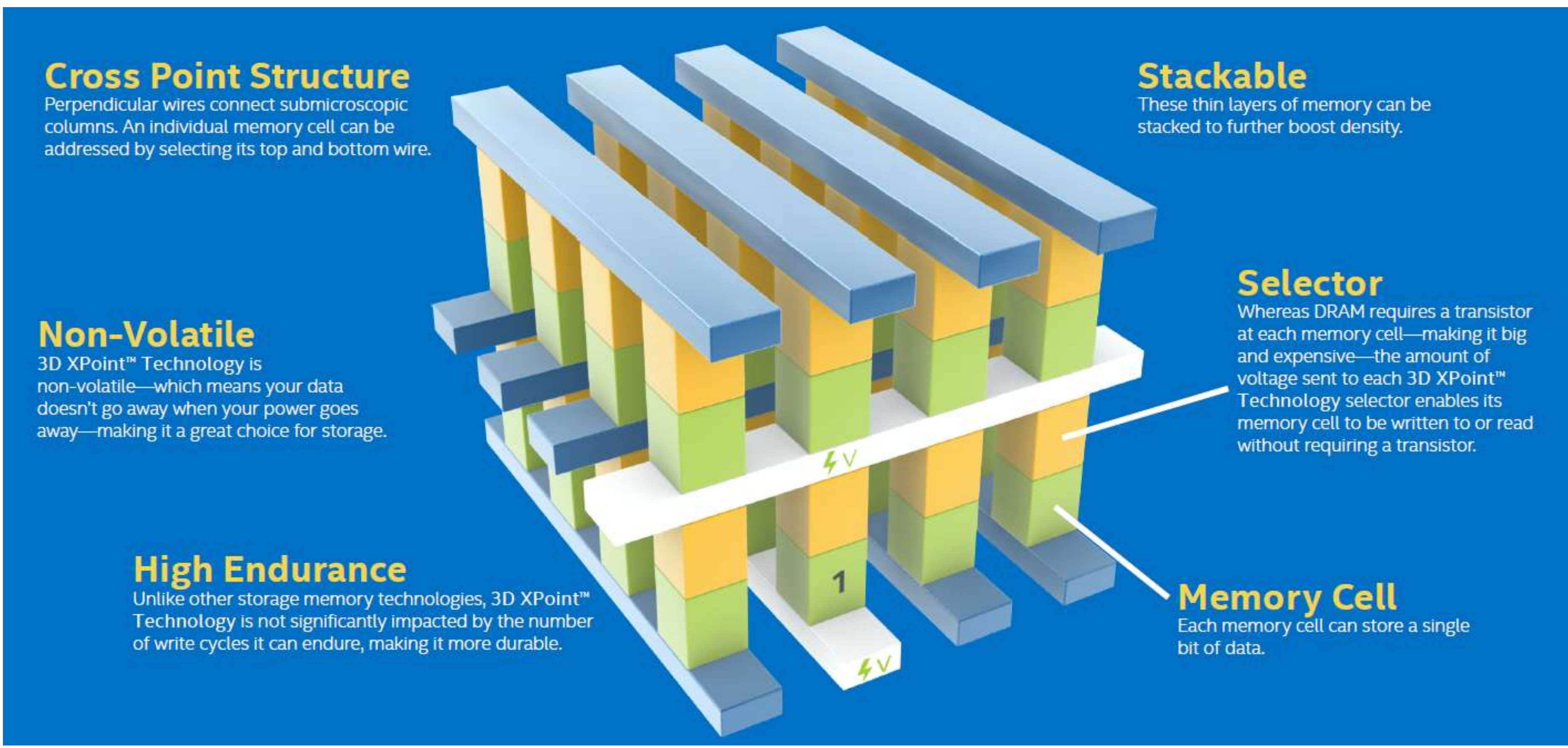
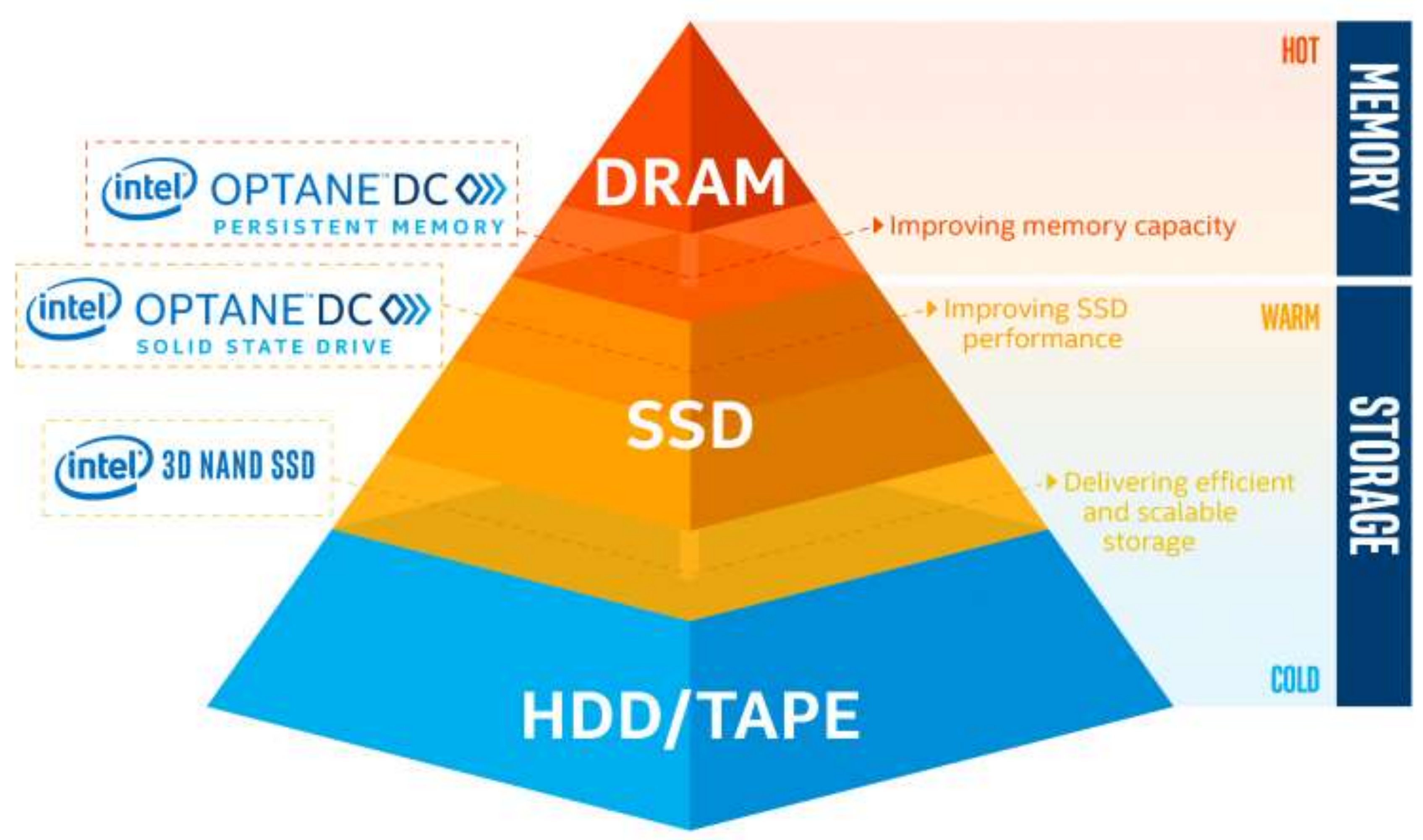


Figure: 3D XPoint memory [225]

Optane memory as a new tier in the memory and storage hierarchy [181]



4.3.2 Key innovations of the Cascade Lake lines (10)

Optane memory – implementation alternatives

Optane memory alternatives

Optane Memory cache [179]
(e.g. 32GB M.2 80MM SSD)

M.2 card

16/32 GB

WD cache in DTs or notebooks

Since Kaby Lake H/S Series
200 Series chipset

Available: 4/2017

Optane SSD DC [178]
(*e.g. P4800X Series cache)*

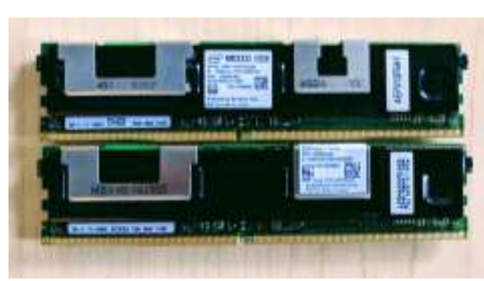
PCIe 3.0 x4 NVMe add-in card

375/750 GB

In datacenters as fast storage or extended memory

Available: 3/2017

Optane DC [180]
Persistent Memory



DDR4 DIMM

128/256/512GB

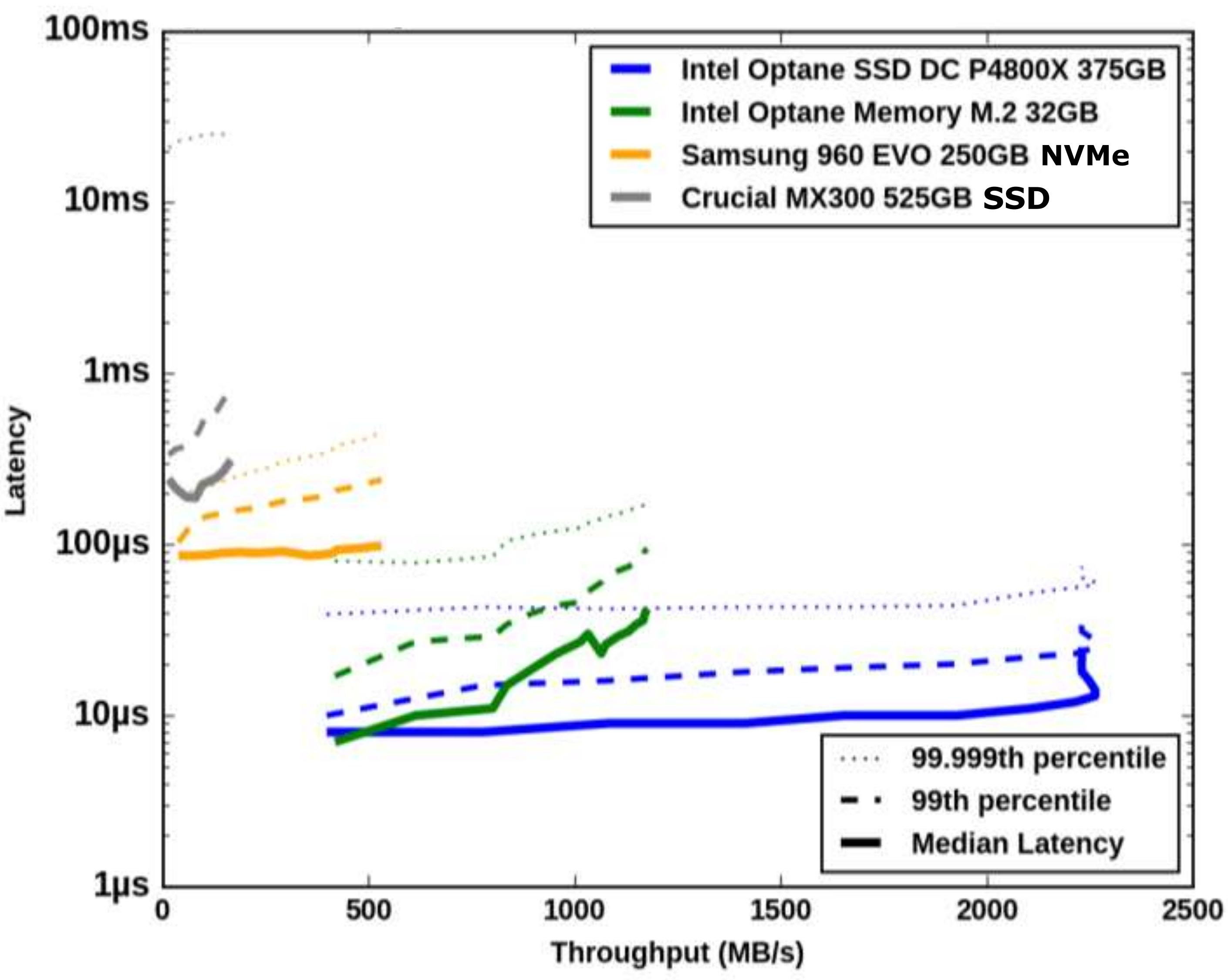
In servers as a new tier of memory

Since: Cascade Lake-AP

Sampled: 05/2018
Broad availability: 2019

Remark: The P4800X Optane SSD DX with the cover plate removed [184]

Random read throughput and latency (Queue Depth: 1.16, 1-4 threads) [183]



Endurance of Optane memory [182]

- There are data available about the endurance of the Optane SSD DC P4800X, as follows.
- At launch Intel stated a write endurance of 30 Drive Writes Per Day (DWPD) for 3 years.
- Later, in 2018, Intel extended their rating for 5 years, based on accumulated reliability data.
- According to industry sources, a further extension to 60 DWPD can be expected [182].

c) Hardware mitigations against security threats [177]

Cascade Lake implements hardware mitigations against targeted side-channel methods

Variant	Side-Channel Method	Mitigation on Cascade Lake
Variant 1	Bounds Check Bypass	OS/VMM
Variant 2	Branch Target Injection	Hardware + OS/VMM
Variant 3	Rogue Data Cache Load	Hardware
Variant 3a	Rogue System Register Read	Firmware
Variant 4	Speculative Store Bypass	Firmware + OS/VMM or runtime
	L1 Terminal Fault	Hardware

Cascade Lake SP expected to provide higher performance over software mitigations available for existing products

For additional information related to security updates and side channel methods on Intel® products, please visit https://www.intel.com/content/www/us/en/architecture-and-technology/facts-about-side-channel-analysis-and-intel-products.html

4.3.2 Key innovations of the Cascade Lake lines (15)

Cascade Lake server product lines

Cascade Lake server product lines

Cascade Lake-SP line	**Cascade Lake-AP line**
Xeon-SP 82/62/52/42/32xx (Platinum/Gold/Silver/Bronze)	*Intel's EX servers*
Up to 28 cores To be shipped in about 1H/2019	Up to 56 cores To be shipped in about 1H/2019
Section 4.3.3	*Section 4.3.4*

4.3.3 The Cascade Lake-SP line

4.3.3 The Cascade Lake-SP line

Key features of the up to 28-core Cascade Lake-SP processor line [177]

Cascade Lake CPU is designed to be compatible with first-gen Intel® Xeon® Scalable platform

- Same core count, cache size, and I/O speeds as first-gen
- Process tuning, frequency push, targeted performance improvements
- Architectural improvements through targeted instruction set enhancements
- New platform capabilities with support for Intel® Optane™ DC persistent memory
- Hardware enhancements for protection against side-channel methods

Features	Cascade Lake CPU
Cores and Threads	Up to 28 Cores and 56 Threads
Last-level Cache	Up to 38.5 MB (non-inclusive)
UPI Speed (GT/s)	Up to 3x UPI @ 10.4 GT/s
PCIe* 3.0 Lanes	Up to 48 lanes with 12 controllers
Memory Speed	Up to 6 channels @ up to 2666 MHz

- Manufactured using the 14++ nm technology.
- To be launched in H1/2019.
- According to industry sources [185] 39 models will be launched with a wide range of core counts, clock speeds and TDP values, as shown below.

Intel Cascade Lake-SP Xeon Platinum Processors [185]

Model	Cores	Base Clock	TDP
Intel Xeon Platinum 8280M	28	2.7 GHz	205W
Intel Xeon Platinum 8280L	28	2.7 GHz	205W
Intel Xeon Platinum 8280	28	2.7 GHz	205W
Intel Xeon Platinum 8276M	28	2.3 GHz	165W
Intel Xeon Platinum 8276L	28	2.3 GHz	165W
Intel Xeon Platinum 8276	28	2.3 GHz	165W
Intel Xeon Platinum 8270	26	2.6 GHz	205W
Intel Xeon Platinum 8268	24	2.9 GHz	205W
Intel Xeon Platinum 8260M	24	2.4 GHz	165W
Intel Xeon Platinum 8260L	24	2.4 GHz	165W
Intel Xeon Platinum 8260	24	2.4 GHz	165W
Intel Xeon Platinum 8260C	24 / 20 / 16	2.4 GHz / 2.6 GHz / 2.8 GHz	65W

Intel Cascade Lake-SP Xeon Gold Processors [185]

Model	Core	Base Clock	TDP
Intel Xeon Gold 6252	24	2.1 GHz	150W
Intel Xeon Gold 6238T	22	2.0 GHz	125W
Intel Xeon Gold 6248	20	2.6 GHz	150W
Intel Xeon Gold 6230	20	2.1 GHz	125W
Intel Xeon Gold 6254	18	3.2 GHz	200W
Intel Xeon Gold 6240	18	2.6 GHz	150W
Intel Xeon Gold 6240C	18 / 14 / 8	2.6 GHz / 2.8 GHz / 3.1 GHz	150W
Intel Xeon Gold 5250	18	2.9 GHz	125W
Intel Xeon Gold 6242	16	2.8 GHz	150W
Intel Xeon Gold 5218	16	2.3 GHz	125W
Intel Xeon Gold 5128T	16	2.2 GHz	105W
Intel Xeon Gold 5117	14	2.0 GHz	105W
Intel Xeon Gold 5215M	10	2.6 GHz	85W
Intel Xeon Gold 5215L	10	2.6 GHz	85W
Intel Xeon Gold 5215	10	2.6 GHz	125W
Intel Xeon Gold 6244	8	3.7 GHz	165W
Intel Xeon Gold 5217M	8	3.0 GHz	125W
Intel Xeon Gold 5217L	8	3.0 GHz	125W
Intel Xeon Gold 5217	8	3.0 GHz	85W

Intel Cascade Lake-SP Xeon Silver Processors [185]

Model	Cores	Base Clock	TDP
Intel Xeon Silver 4216	16	2.2 GHz	100W
Intel Xeon Silver 4214	12	2.2 GHz	85W
Intel Xeon Silver 4214C	12 / 10 /8	2.1 GHz /2.2 GHz / 2.3GHz	105W
Intel Xeon Silver 4210	10	2.2 GHz	85W
Intel Xeon Silver 4215	8	2.5 GHz	85W
Intel Xeon Silver 4209T	8	2.2 GHz	70W

Intel Cascade Lake-SP Xeon Bronze Processors [185]

Model	Cores	Base Clock	TDP
Intel Xeon Bronze 3204	6	1.9 GHz	85W

4.3.4 The Cascade Lake-AP line

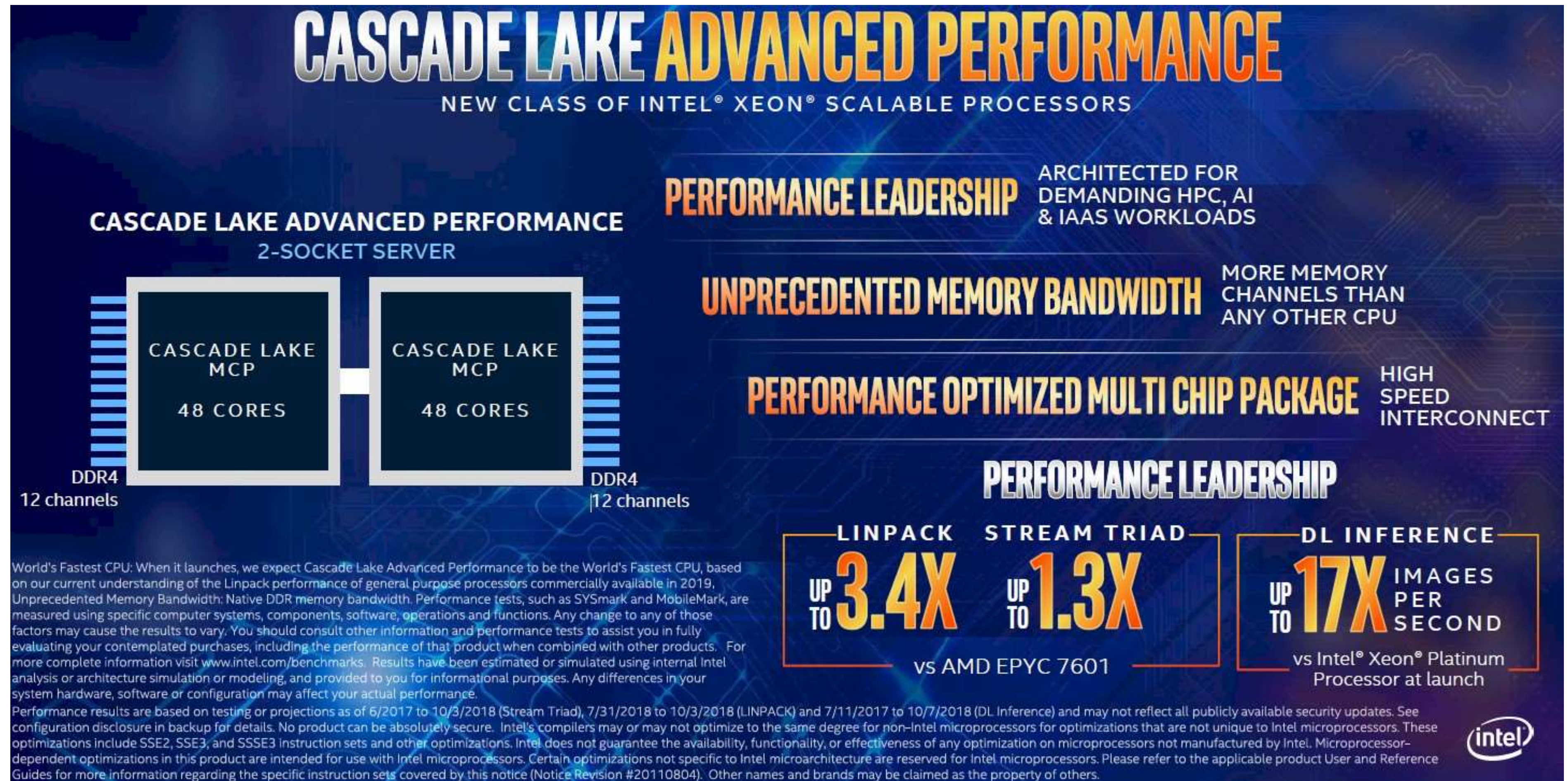
4.3.4 The Cascade Lake-AP line

- AP means Advanced Processor.
- The Cascade Lake-AP line has 48 cores and is targeting
  - demanding high-performance computing (HPC),
  - artificial intelligence (AI/DL) and
  - infrastructure-as-a-service (IaaS)

  workloads.
- The Cascade Lake-AP is manufactured on 14++ nm technology, and is an MCM design including two dies, each with 24 cores.

  The roadmap shows a shipping date of Q4/2018 but subsequent statements indicate a later shipping date of H1/2019.
- The 10 nm Ice Lake line is only scheduled for 2020.

Layout and performance of the 48-core Cascade Lake AP [177]



5. References

5. References (1)

[1]: Radhakrisnan S., Sundaram C. and Cheng K., „The Blackford Northbridge Chipset for the Intel 5000", IEEE Micro, March/April 2007, pp. 22-33

[2]: Next-Generation AMD Opteron Processor with Direct Connect Architecture – 4P Server Comparison, http://www.amd.com/us-en/assets/content_type/DownloadableAssets/4P_Server_Comparison_PID_41461.pdf

[3]: Intel® 5000P/5000V/5000Z Chipset Memory Controller Hub (MCH) – Datasheet, Sept. 2006. http://www.intel.com/design/chipsets/datashts/313071.htm

[4]: Intel® E8501 Chipset North Bridge (NB) Datasheet, Mai 2006, http://www.intel.com/design/chipsets/e8501/datashts/309620.htm

[5]: Conway P & Hughes B., „The AMD Opteron Northbridge Architecture", IEEE MICRO, March/April 2007, pp. 10-21

[6]: Intel® 7300 Chipset Memory Controller Hub (MCH) – Datasheet, Sept. 2007, http://www.intel.com/design/chipsets/datashts/313082.htm

[7]: Supermicro Motherboards, http://www.supermicro.com/products/motherboard/

[8]: Sander B., „AMD Microprocessor Technologies," 2006, http://www.ewh.ieee.org/r4/chicago/foxvalley/IEEE_AMD_Meeting.ppt

[9]: AMD Quad FX Platform with Dual Socket Direct Connect (DSDC) Architecture, http://www.asisupport.com/ts_amd_quad_fx.htm

[10]: Asustek motherboards - http://www.asus.com.tw/products.aspx?l1=9&l2=39
http://support.asus.com/download/model_list.aspx?product=5&SLanguage=en-us

5. References (2)

[11]: Kanter, D. „A Preview of Intel's Bensley Platform (Part I)," Real Word Technologies, Aug. 2005, http://www.realworldtech.com/page.cfm?ArticleID=RWT110805135916&p=2

[12]: Kanter, D. „A Preview of Intel's Bensley Platform (Part II)," Real Word Technologies, Nov. 2005, http://www.realworldtech.com/page.cfm?ArticleID=RWT112905011743&p=7

[13]: Quad-Core Intel® Xeon® Processor 7300 Series Product Brief, Intel, Nov. 2007 http://download.intel.com/products/processor/xeon/7300_prodbrief.pdf

[14]: „AMD Shows Off More Quad-Core Server Processors Benchmark" X-bit labs, Nov. 2007 http://www.xbitlabs.com/news/cpu/display/20070702235635.html

[15]: AMD, Nov. 2006 http://www.asisupport.com/ts_amd_quad_fx.htm

[16]: Rusu S., "A Dual-Core Multi-Threaded Xeon Processor with 16 MB L3 Cache," Intel, 2006, http://ewh.ieee.org/r5/denver/sscs/Presentations/2006_04_Rusu.pdf

[17]: Goto H., Intel Processors, PCWatch, March 04 2005, http://pc.watch.impress.co.jp/docs/2005/0304/kaigai162.htm

[18]: Gilbert J. D., Hunt S., Gunadi D., Srinivas G., "The Tulsa Processor," Hot Chips 18, 2006, http://www.hotchips.org/archives/hc18/3_Tues/HC18.S9/HC18.S9T1.pdf

[19]: Goto H., IDF 2007 Spring, PC Watch, April 26 2007, http://pc.watch.impress.co.jp/docs/2007/0426/hot481.htm

[20]: Hruska J., "Details slip on upcoming Intel Dunnington six-core processor," Ars technica, February 26, 2008, http://arstechnica.com/news.ars/post/20080226-details-slip-on-upcoming-intel-dunnington-six-core-processor.html

[21]: Goto H., 32 nm Westmere arrives in 2009-2010, PC Watch, March 26 2008, http://pc.watch.impress.co.jp/docs/2008/0326/kaigai428.htm

[22]: Singhal R., "Next Generation Intel Microarchitecture (Nehalem) Family: Architecture Insight and Power Management, IDF Taipeh, Oct. 2008, http://intel.wingateweb.com/taiwan08/published/sessions/TPTS001/FA08%20IDF-Taipei_TPTS001_100.pdf

[23]: Smith S. L., "45 nm Product Press Briefing,", IDF Fall 2007, ftp://download.intel.com/pressroom/kits/events/idffall_2007/BriefingSmith45nm.pdf

[24]: Bryant D., "Intel Hitting on All Cylinders," UBS Conf., Nov. 2007, http://files.shareholder.com/downloads/INTC/0x0x191011/e2b3bcc5-0a37-4d06-aa5a-0c46e8a1a76d/UBSConfNov2007Bryant.pdf

[25]: Barcelona's Innovative Architecture Is Driven by a New Shared Cache, http://developer.amd.com/documentation/articles/pages/8142007173.aspx

[26]: Larger L3 cache in Shanghai, Nov. 13 2008, AMD, http://forums.amd.com/devblog/blogpost.cfm?threadid=103010&catid=271

[27]: Shimpi A. L., "Barcelona Architecture: AMD on the Counterattack," March 1 2007, Anandtech, http://www.anandtech.com/cpuchipsets/showdoc.aspx?i=2939&p=1

5. References (4)

[28]: Rivas M., “Roadmap update,”, 2007 Financial Analyst Day, Dec. 2007, AMD, http://download.amd.com/Corporate/MarioRivasDec2007AMDAnalystDay.pdf

[29]: Scansen D., “Under the Hood: AMD’s Shanghai marks move to 45 nm node,” EE Times, Nov. 11 2008, http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=212002243

[30]: 2-way Intel Dempsey/Woodcrest CPU Bensley Server Platform, Tyan, http://www.tyan.com/tempest/training/s5370.pdf

[31]: Gelsinger P. P., “Intel Architecture Press Briefing,”, 17. March 2008, http://download.intel.com/pressroom/archive/reference/Gelsinger_briefing_0308.pdf

[32]: Mueller S., Soper M. E., Sosinsky B., Server Chipsets, Jun 12, 2006, http://www.informit.com/articles/article.aspx?p=481869

[33]: Goto H., IDF, Aug. 26 2005, http://pc.watch.impress.co.jp/docs/2005/0826/kaigai207.htm

[34]: TechChannel, http://www.tecchannel.de/_misc/img/detail1000.cfm?pk=342850&fk=432919&id=il-74145482909021379

[35]: Intel quadcore Xeon 5300 review, Nov. 13 2006, Hardware.Info, http://www.hardware.info/en-US/articles/amdnY2ppZGWa/Intel_quadcore_Xeon_5300_review

[36]: Wasson S., Intel's Woodcrest processor previewed, The Bensley server platform debuts, Mai 23, 2006, The Tech Report, http://techreport.com/articles.x/10021/1

5. References (5)

[37]: Enderle R., AMD Shanghai "We are back! TGDaily, November 13, 2008, http://www.tgdaily.com/content/view/40176/128/

[38]: Clark J. & Whitehead R., "AMD Shanghai Launch, Anandtech, Nov. 13 2008, http://www.anandtech.com/showdoc.aspx?i=3456

[39]: Chiappetta M., AMD Barcelona Architecture Launch: Native Quad-Core, Hothardware, Sept. 10, 2007, http://hothardware.com/Articles/AMD_Barcelona_Architecture_Launch_Native_QuadCore/

[40]: Hachman M., "AMD Phenom, Barcelona Chips Hit By Lock-up Bug,", ExtremeTech, Dec. 5 2007, http://www.extremetech.com/article2/0,2845,2228878,00.asp

[41]: AMD Opteron™ Processor for Servers and Workstations, http://amd.com.cn/CHCN/Processors/ProductInformation/0,,30_118_8826_8832,00-1.html

[42]: AMD Opteron Processor with Direct Connect Architecture, 2P Server Power Savings Comparison, AMD, http://enterprise.amd.com/downloads/2P_Power_PID_41497.pdf

[43]: AMD Opteron Processor with Direct Connect Architecture, 4P Server Power Savings Comparison, AMD, http://enterprise.amd.com/downloads/4P_Power_PID_41498.pdf

[44]: AMD Opteron Product Data Sheet, AMD, http://pdfs.icecat.biz/pdf/1868812-2278.pdf

[45]: Images, Xtreview,
http://xtreview.com/images/K10%20processor%2045nm%20architec%203.jpg

[46]: Kanter D., "Inside Barcelona: AMD's Next Generation, Real World Tech., Mai 16 2007,
http://www.realworldtech.com/page.cfm?ArticleID=RWT051607033728

[47]: Kanter D,, "AMD's K8L and 4x4 Preview, Real World Tech. June 02 2006,
http://www.realworldtech.com/page.cfm?ArticleID=RWT060206035626&p=1

[48]: Kottapali S., Baxter J., "Nehalem-EX CPU Architecture", Hot Chips 21, 2009,
http://www.hotchips.org/archives/hc21/2_mon/HC21.24.100.ServerSystemsI-Epub/HC21.24.122-Kottapalli-Intel-NHM-EX.pdf

[49]: Kahn O., Piazza T., Valentine B., "Technology Insight: Intel next Generation Microarchitecture Codename Sandy Bridge, IDF, San Francisco, Sept. 2010

[50]: Piazza T., Jiang H., "Intel Next Generation Microarchitecture Codename Sandy Bridge: Processor Graphics, IDF, San Francisco, Sept. 2010

[51]: Braun-Protte T., "Intel die neuste Generation", March 2010,
https://sp.ts.fujitsu.com/dmsp/docs/02_fujitsu_intel_server_cpu.pdf

[52]: De Gelas J., The Intel Xeon E7-8800 v3 Review: The POWER8 Killer?, AnandTech, May 8 2015, http://www.anandtech.com/show/9193/the-xeon-e78800-v3-review

[53]: Tyan Confidential, Tempest i5000VF S5370, 2-way Intel Dempsey/Woodcrest CPU Bensley Server Platform, http://www.tyan.com/tempest/training/s5370.pdf

[54]: Intel 5520 Chipset and Intel 5500 Chipset, Datasheet, March 2009, http://www.intel.com/content/www/us/en/chipsets/5520-5500-chipset-ioh-datasheet.html

[55]: Neiger G., Santoni A., Leung F., Rodgers D., Uhlig R., Intel® Virtualization Technology: Hardware support for efficient processor virtualization, Aug. 10 2006, Vol. 10, Issue 3, http://www.intel.com/technology/itj/2006/v10i3/1-hardware/1-abstract.htm

[56]: Intel Software Networks: Forums, http://software.intel.com/en-us/forums/showthread.php?t=56802

[57]: Wikipedia: x86 virtualization, 2011, http://en.wikipedia.org/wiki/X86_virtualization#Intel_virtualization_.28VT-x.29

[58]: Sharma D. D., Intel 5520 Chipset: An I/O Hub Chipset for Server, Workstation, and High End Desktop, Hotchips 2009, http://www.hotchips.org/archives/hc21/2_mon/HC21.24.200.I-O-Epub/HC21.24.230.DasSharma-Intel-5520-Chipset.pdf

[59]: Morgan T. P., Intel's future Sandy Bridge Xeons exposed, May 30 2011, The Register, http://www.theregister.co.uk/2011/05/30/intel_sandy_bridge_xeon_platforms/page2.html

[60]: Gilbert J. D., Hunt S. H., Gunadi D., Srinivas G., The Tulsa Processor: A Dual Core Large Shared-Cache Intel Xeon Processor 7000 Sequence for the MP Server Market Segment, Aug 21 2006, http://www.hotchips.org/archives/hc18/3_Tues/HC18.S9/HC18.S9T1.pdf

[61]: Intel Server Board Set SE8500HW4, Technical Product Specification, Revision 1.0, May 2005, ftp://download.intel.com/support/motherboards/server/sb/se8500hw4_board_set_tpsr10.pdf

[62]: Mitchell D., Intel Nehalem-EX review, PCPro,
http://www.pcpro.co.uk/reviews/processors/357709/intel-nehalem-ex

[63]: Nagaraj D., Kottapalli S.: Westmere-EX: A 20 thread server CPU, Hot Chips 2010
http://www.hotchips.org/uploads/archive22/HC22.24.610-Nagara-Intel-6-Westmere-EX.pdf

[64]: Intel Xeon Processor E7-8800/4800/2800 Product Families, Datasheet Vol. 1 of 2, April 2011, http://www.intel.com/Assets/PDF/datasheet/325119.pdf

[65]: Supermicro P4QH6 / P4QH8 User's Manual, 2002,
http://www.supermicro.com/manuals/motherboard/GC-HE/MNL-0665.pdf

[66]: Intel Xeon Processor 7500/6500 Series, Public Gold Presentation, March 30 2010,
http://cache-www.intel.com/cd/00/00/44/64/446456_446456.pdf

[67]: Supermicro X8QB6-F / X8QBE-F User's Manual, 2010,
http://files.siliconmechanics.com/Documentation/Rackform/iServ/R413/Mainboard/MNL-X8QB-E-6-F.pdf

[68]: Rusu & al.: A 45 nm 8-Core Enterprise Xeon Processor, IEEE journal of Solid State Circuits, Vol. 45, No. 1, Jan. 2010, pp. 7-14

[69]: Jafarjead B., "Intel Core Duo Processor," Intel, 2006,
http://masih0111.persiangig.com/document/peresentation/behrooz%20jafarnejad.ppt

[70]: Keutzer K., Malik S., Newton R., Rabaey J., Sangiovanni-Vincentelli A., System Level Design: Orthogonalization of Concerns and Platform-Based Design, IEEE Transactions on Computer-Aided Design of Circuits and Systems, Vol. 19, No. 12, Dec. 2000, pp.?????

[71]: Perich D., Intel Volume platforms Technology Leadership, Presentation at HP World 2004, http://98.190.245.141:8080/Proceed/HPW04CD/papers/4194.pdf

[72]: Krazit T., Intel Sheds Light on 2005 Desktop Strategy, IDG News Service, Dec. 07 2004, http://pcworld.about.net/news/Dec072004id118866.htm

[73]: Molka D., Hackenberg D., Schöne R., Müller M. S., Memory Performance and Cache Coherency Effects on an Intel Nehalem multiprocessor System, Proc. 18th Int. Conf. on Parallel Architectures and Compilation, Techniques, PACT 2009, http://doi.ieeecomputersociety.org/10.1109/PACT.2009.22

[74]: Radhakrishnan S., Chinthamani S., Cheng K., The Blackford Northbridge Chipset for Intel 5000, IEEE MICRO, March-April 2007, pp. 22-33

[75]: Rolf T., Cache Organization and Memory Management of the Intel Nehalem Computer Architecture, University of Utah, Dec. 2009, http://rolfed.com/nehalem/nehalemPaper.pdf

[76]: Levinthal D., Performance Analysis Guide for Intel Core i7 Processor and Intel Xeon 5500 processors, Version 1.0, 2008 http://software.intel.com/sites/products/collateral/hpc/vtune/performance_analysis_guide.pdf

[77]: Ryszard Sommefeldt R., Intel Xeon 3.4GHz ['Nocona' core], 18th Aug, 2004, http://www.hexus.net/content/item.php?item=822&page=1&search=reas

[78]: 3.6GHz 2MB 800MHz Intel Xeon Processor SL7ZC C8511, http://store.flagshiptech.com/products/3.6GHz-2MB-800MHz-Intel-Xeon-Processor-SL7ZC-C8511.html

[79]: Gavrichenkov I., Workstation Processors Duel: AMD Opteron against Intel Xeon, Page 5 12/21/2005, http://www.xbitlabs.com/articles/cpu/display/opteron-xeon-workstation_5.html

[80]: Glaskowsky P.: Investigating Intel's Lynnfield mysteries, cnet News, Sept. 21. 2009, http://news.cnet.com/8301-13512_3-10357328-23.html

[81]: Kurd & al., Westmere: A Family of 32 nm IA Processors, Proc. ISSCC 2010, pp. 96-98

[82]: Miller M. J.,Intel Previews ISSCC: 6-Core Gulftown Processor, Feb 03, 2010, http://forwardthinking.pcmag.com/chips/282694-intel-previews-isscc-6-core-gulftown-processor

[83]: Hill D., Chowdhury M., Westmere Xeon-56xx "Tick" CPU, Hot Chips 22, 2010, http://www.hotchips.org/uploads/archive22/HC22.24.620-Hill-Intel-WSM-EP-print.pdf

[84]: Pawlowski S.: Intelligent and Expandable High- End Intel Server Platform, Codenamed Nehalem-EX, IDF 2009

[85]: Intel Sandy Bridge Review, Bit-tech, Jan. 3 2011, http://www.bit-tech.net/hardware/cpus/2011/01/03/intel-sandy-bridge-review/1

[86]: Kahn O., Piazza T., Valentine B.: Technology Insight: Intel Next Generation Microarchitecture Codename Sandy Bridge, IDF 2010 extreme.pcgameshardware.de/.../281270d1288260884-bonusmaterial-pc- games-ardware-12-2010-sf10_spcs001_100.pdf

[87]: Inkley B., Tetrick S., Intel Multi-core Architecture and Implementations, IDF, Session MATS002, March 7 2006

[88]: Smith S. L., Intel Roadmap Overview, IDF Fall, Aug. 20 2008 http://download.intel.com/pressroom/kits/events/idffall_2008/SSmith_briefing_roadmap.p

[89]: Intel Public Roadmap; Business Platforms for Desktop, Mobile & Data Center, H1' 2011 http://www.2ncinesil.com/download/Public%20Roadmap%201H%202011.pptx

[90]: Rusu S., „Circuit Technologies for Multi-Core Processor Design," April 2006, http://www.ewh.ieee.org/r6/scv/ssc/April06.pdf

[91]: Goto H., IDF, Aug. 26 2005, http://pc.watch.impress.co.jp/docs/2005/0826/kaigai207.htm

[92]: TechChannel, http://www.tecchannel.de/_misc/img/detail1000.cfm?pk=342850&fk=432919&id=il-7414548290902137 9

[93]: Sandy Bridge, http://en.wikipedia.org/wiki/Sandy_Bridge

[94]: Morgan T. P., Intel's future Sandy Bridge Xeons exposed, The Register, May 30. 2011, http://www.theregister.co.uk/2011/05/30/intel_sandy_bridge_xeon_platforms

[95]: Gelsinger P., Keynote, IDF Spring 2005, March 2005, http://www.intel.com/pressroom/kits/events/idfspr_2005/index.htm

[96]: Kuppuswamy, R. C al., Over one million TPCC with a 45nm 6-core Xeon® CPU, IEEE International Solid-State Circuits Conference - Digest of Technical Papers, ISSCC 2009, Feb. 2009, pp. 70 - 71,71a

[97]: Perich D., Intel Volume Platforms Technology Leadership, Solutions and Technology Conference, HP World 2004, http://www.openmpe.com/cslproceed/HPW04CD/papers/4194.pdf

[98]: De Gelas J., Westmere-EX: Intel Improves their Xeon Flagship, AnandTech, April 6 2011, http://www.anandtech.com/show/4259/westmereex-intels-flagship-improves

[99]: Starke W.J., POWER7: IBM's Next Generation, Balanced POWER Server Chip, Hot Chips 21, http://www.hotchips.org/wp-content/uploads/hc_archives/hc21/3_tues/HC21.25.800.ServerSystemsII-Epub/HC21.25.835.Starke-IBM-POWER7SystemBalancev13_display.pdf

5. References (12)

[100]: Intel E8500 Chipset eXternal Memory Bridge (XMB), Datasheet, March 2005, http://www.intel.co.id/content/dam/doc/datasheet/e8500-chipset-external-memory-bridge-datasheet.pdf

[101]: Intel Server System SSH4 Board Set, Technical Product Specification, Oct. 2003, http://download.intel.com/support/motherboards/server/ssh4/ssh4_tps.pdf

[102]: Haas J. & Vogt P., Fully buffered DIMM Technology Moves Enterprise Platforms to the Next Level, Technology Intel Magazine, March 2005, pp. 1-7

[103]: Ganesh B., Jaleel A., Wang D., Jacob B., Fully-Buffered DIMM Memory Architectures: Understanding Mechanisms, Overheads and Scaling, 2007, https://ece.umd.edu/~blj/papers/hpca2007.pdf

[104]: Smith S. L., Quad-Core Press Briefing, IDF 2006, http://www.intel.com/pressroom/kits/quadcore/qc_pressbriefing.pdf

[105]: Smith S.L., Intel Enterprise Roadmap, IDF 2010

[106]: Intel 7500 Chipset, Datasheet, March 2010, http://www.intel.com/content/dam/doc/datasheet/7500-chipset-datasheet.pdf

[107]: 7500/7510/7512 Scalable Memory Buffer, Datasheet, April 2011, http://www.intel.com/content/www/us/en/chipsets/7500-7510-7512-scalable-memory-buffer-datasheet.html

[108]: X8QB6-LF User Manual, Supermicro

[109]: Baseboard management controller (BMC) definition, TechTarget, May 2007, http://searchnetworking.techtarget.com/definition/baseboard-management-controller

[110]: Morgan T. P., Intel (finally) uncages Nehalem-EX beast, The Register, March 30 2010, http://www.theregister.co.uk/2010/03/30/intel_nehalem_ex_launch/?page=2

[111]: Morgan T. P., Intel Puts More Compute Behind Xeon E7 Big Memory, The Platform, May 5 2015, http://www.theplatform.net/2015/05/05/intel-puts-more-compute-behind-xeon-e7-big-memory/

[112]: Niederste-Berg M., 28 core Skylake EP/EX plattform comes into view, Hardware LUXX, 26 May 2015, http://www.hardwareluxx.com/index.php/news/hardware/cpu/35474-28-core-skylake-epex-plattform-comes-into-view.html

[113]: Intel C102/C104 Scalable Memory Buffer, Datasheet, Febr. 2014, http://www.intel.com/content/dam/www/public/us/en/documents/datasheets/c102-c104-scalable-memory-buffer-datasheet.pdf

[114]: Intel Xeon Processor E7-8800/4800 V3 Product Family, Technical Overview, May 21 2015, https://software.intel.com/en-us/articles/intel-xeon-processor-e7-88004800-v3-product-family-technical-overview

[115]: X10QBi Platform with X10QBi Baseboard, User's manual, Supermicro, 2014, ftp://ftp.supermicro.com/CDR-X10-Q_1.00_for_Intel_X10_Q_platform/MANUALS/X10QBi.pdf

[116]: Rusu S., Muljono H., Ayers D., Tam S., A 22nm 15-core enterprise Xeon processor family, IEEE ISSC Journal, Vol. 50, Issue 1, Jan. 2015

[117]: Esmer I., Ivybridge Server Architecture: A Converged Server, Aug. 2014, Hot Chips 2014, http://www.hotchips.org/wp-content/uploads/hc_archives/hc26/HC26-12-day2-epub/HC26.12-8-Big-Iron-Servers-epub/HC26.13.832-IvyBridge-Esmer-Intel-IVB%20Server%20Hotchips2014.pdf

[118]: Chiappetta M., Intel Launches Xeon E7-8800 and E7-4800 v3 Processor Families, Hot Hardware, May 5 2015, http://hothardware.com/reviews/intel-launches-xeon-e7-v3-family-of-processors

[119]: Broadwell-EX/EP disappear from the latest Intel roadmap, AnandTech Forums, http://forums.anandtech.com/showthread.php?t=2439161

[120]: Accelerate Big Data Insights with the Intel Xeon Processor E7-8800/4800 v3 Product Families, Product Brief, 2015, http://www.intel.com/newsroom/kits/xeon/e7v3/pdfs/Xeon_E7v3_ProductBrief.pdf

[121]: Accelerating Silicon Design with the Intel Xeon Processor E7-8800 v3 Product Family, 2015, http://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/xeon-processor-e7-8800-v3-brief.pdf

[122]: Broadwell-EX/EP disappeared from Intel's enterprise roadmaps, Linus Tech Tips, July 16 2015, http://linustechtips.com/main/topic/409185-broadwell-exep-disappeared-from-intels-enterprise-roadmaps/

[123]: van Monsjou D., Massive Leak shows details on Skylake Xeon Chips, Overclock3D, May 25 2015, http://www.overclock3d.net/articles/cpu_mainboard/massive_leak_shows_details_on_skylake_xeon_chips/1

[124]: Server duel: Xeon Woodcrest vs. Opteron Socket F, Tweakers, 7 September 2006, http://tweakers.net/reviews/646/2/server-duel-xeon-woodcrest-vs-opteron-socket-f-post-mortem-netburst.html

[125]: Papazian I. E. & al., Ivy Bridge Server: A Converged Design, IEEE MICRO, March-April 2015, pp. 16-25, http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=7091791

[126]: Gao Y., HPC Workload Performance Tuning on POWER8 with IBM XL Compilers and Libraries, SPXXL/Scicomp Summer Workshop 2014, http://spscicomp.org/wordpress/wp-content/uploads/2014/05/gao-IBM-XL-compilers-for-POWER8-Scicomp-2014.pdf

[127]: New Xeon Processor Numbering, More Than Just a Number, Intel.com, April 5 2011, https://communities.intel.com/community/itpeernetwork/datastack/blog/2011/04/05/new-xeon-processor-numbering-more-than-just-a-number

[128]: Pirzada U., Intel Launches Its 14nm Broadwell-EX Platform,WCCF Tech, 07 June 2016, http://wccftech.com/intel-broadwell-ex-xeon-e7-8890-24-cores/

[129]: Karedla, R., Intel Xeon E5-2600 v3 (Haswell) Architecture & Features, Intel, 2014, http://repnop.org/pd/slides/PD_Haswell_Architecture.pdf

[130]: Syamalakumari S., Intel® Xeon® Processor E7-8800/4800 V3 Product Family Technical Overview, Intel, May 21 2015, https://software.intel.com/en-us/articles/intel-xeon-processor-e7-88004800-v3-product-family-technical-overview

[131]: Rotem E. et al., Multiple Clock and Voltage Domains for Chip Multi Processors, Proc. 42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO) December 12–16, 2009, New York, NY, USA.

[132]: Gianos C., Intel Xeon Processor E5-2600 v3 Product Family Architectural Overview, Intel, Nov. 16, 2014, https://www.yumpu.com/en/document/view/34127638/intelr-xeonr-processor-e5-2600-v3-overview-for-sc14

[133]: US 2014/0229750 A1 Patent application Method and apparatus for per core performance states, Filed March 13 2012, Published Aug. 14 2014

[134]: Klemm M., Programming and Tuning for Intel® Xeon® Processors, Intel, 2015, https://www.dkrz.de/Nutzerportal-en/doku/training/introduction-to-mistral-july-2014/Programming_and_Tuning_for_Intel_Xeon_Processors_2015-07-01_M.Klemm.pdf?lang=en

[135]: https://en.wikipedia.org/wiki/RC_circuit

[136]: Samsung Memory DDR4 SDRAM, Broshure, Samsung, 2015, http://www.samsung.com/semiconductor/global/file/insight/2015/11/DDR4_Brochure_Nov-15-0.pdf

[137]: Hibben M., Will IBM Sell Its Mainframe Business? Seeking Alpha, Jul. 20, 2016 http://seekingalpha.com/article/3989959-will-ibm-sell-mainframe-business?auth_param=djq9a:1bovpmu:2be6695ecca2e22ac60039ce2a123ea7

[138]: Intel QuickPath Interconnect Electrical Architecture Overview, Intel, 2010, http://jamesstovalldesign.files.wordpress.com/2012/08/pages-from-intel_qpied_final_03-18-10.pdf

[139]: Kumar A., Trivedi M., Intel Xeon Scalable Processor Architecture Deep Dive, June 12 2017, https://www.primeline-solutions.de/files/intel-xeon-scalable-architecture-deep-dive_1.pdf

5. References (17)

[140]: De Gelas J., Cutress I., Sizing Up Servers: Intel's Skylake-SP Xeon versus AMD's EPYC 7000 - The Server CPU Battle of the Decade?, AnandTech, July 11 2017, http://www.anandtech.com/show/11544/intel-skylake-ep-vs-amd-epyc-7000-cpu-battle-of-the-decade/7

[141]: De Gelas J., The Intel Xeon E5 v4 Review: Testing Broadwell-EP With Demanding Server Workloads, AnandTech, March 31 2016, http://www.anandtech.com/print/10158/the-intel-xeon-e5-v4-review

[142]: Intel Xeon Processor E5-1600/E5-2600/E5-4600 Product Families, Datasheet, Vol. One, May 2012, https://www.intel.com/content/dam/www/public/us/en/documents/datasheets/xeon-e5-1600-2600-vol-1-datasheet.pdf

[143]: Intel Xeon Scalable Platform, Newsroom, 2017, https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/07/intel-xeon-scalable-processors-overview.pdf

[144]: Kennedy P., Intel Xeon Scalable Processor Family Platform Level Overview, July 11 2017, STH, https://www.servethehome.com/intel-xeon-scalable-processor-family-platform-level-overview/

[145]: Shrater J., Intel's new innovation Engine Enables Differentiatied Firmware, Aug. 19 2015, https://embedded.communities.intel.com/thread/12385

[146]: Shrout R., Intel Xeon Scalable Processor Launch - New Architecture, New Platform for Data Center, PC Perspective, July 11 2017, https://www.pcper.com/reviews/Processors/Intel-Xeon-Scalable-Processor-Launch-New-Architecture-New-Platform-Data-Center

[147]: Chiappetta M., Intel Xeon Scalable Debuts: Dual Xeon Platinum 8176 With 112 Threads Tested, Hot Hardware, July 11 2017, https://hothardware.com/reviews/intel-xeon-scalable-processor-family-review

[148]: Hruska J., AMD Epyc Faces Off With Intel Skylake-SP Xeon in Massive Server Battle, ExtremeTech, July 12 2017, https://www.extremetech.com/computing/252262-amd-epyc-faces-off-intel-skylake-sp-xeon-massive-server-battle

[149]: Intel Xeon Scalable Platform, Product Brief, Intel, July 2017, https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/07/intel-xeon-scalable-processors-product-brief.pdf

[150]: Safavi M., Thermal Design And NEBS Compliance, Electronics Cooling, Febr. 1 2006, https://www.electronics-cooling.com/2006/02/thermal-design-and-nebs-compliance/

[151]: Morgan T. P., The X86 Battle Lines Drawn With Intel's Skylake Launch, The Next Platform, July 11 2017, https://www.nextplatform.com/2017/07/11/x86-battle-lines-drawn-intels-skylake-launch/

[152]: Alcorn P., Intel Introduces New Mesh Architecture For Xeon And Skylake-X Processors, Tom's Hardware, June 15 2017, http://www.tomshardware.com/news/intel-mesh-architecture-skylake-x-hedt,34806.html

[153]: Morgan T. P., Big Iron Xeons Get A Broadwell Compute And Memory Boost, The Next Platform, June 6 2016, https://www.nextplatform.com/2016/06/06/big-iron-xeons-get-broadwell-compute-memory-boost/

[154]: Mulnix D., Intel Xeon Processor Scalable Family Technical Overview, June 22 2017, https://software.intel.com/en-us/articles/intel-xeon-processor-scalable-family-technical-overview

[155]: Cutress I., The Intel Skylake-X Review: Core i9 7900X, i7 7820X and i7 7800X Tested, AnandTech, June 19 2017, http://www.anandtech.com/print/11550/the-intel-skylakex-review-core-i9-7900x-i7-7820x-and-i7-7800x-tested

[156]: Intel 64 and IA-32 Architectures Optimization Reference Manual, June 2016, https://www.intel.com/content/dam/www/public/us/en/documents/manuals/64-ia-32-architectures-optimization-manual.pdf

[157]: Intel Xeon Phi Coprocessor Datasheet, Nov. 2012, http://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/xeon-phi-datasheet.pdf

[158]: Deploying HPC Cluster with Mellanox InfiniBand Interconnect Solutions, Reference Design, June 2014, http://www.mellanox.com/related-docs/solutions/deploying-hpc-cluster-with-mellanox-infiniband-interconnect-solutions.pdf

[159]: Wardrope I., High Performance Computing - Driving Innovation and Capability, 2013, http://www2.epcc.ed.ac.uk/downloads/lectures/IanWardrope.pdf

[160]: QLogic TrueScale InfiniBand, the Real Value, Solutions for High Performance Computing, Technology Brief, 2009, http://www.transtec.de/fileadmin/Medien/pdf/HPC/qlogic/qlogic_techbrief_truescale.pdf

[161]: Intel Scalable System Framework Intel HPC Developer Conference, Nov. 12-13 2016, http://itoc.sjtu.edu.cn/wp-content/uploads/2016/05/Intel-Omni-Path-in-HPC.pdf

[162]: Ng P. K., High-End Desktop Platform Design Overview for the Next Generation Intel Microarchitecture (Nehalem) Processor, IDF Taipei, 2008, TDPS001, Intel, http://intel.wingateweb.com/taiwan08/published/sessions/TDPS001/FA08%20IDF-Taipei_TDPS001_100.pdf

[163]: [63]: Nagaraj D., Kottapalli S.: Westmere-EX: A 20 thread server CPU, Hot Chips 2010 http://www.hotchips.org/uploads/archive22/HC22.24.610-Nagara-Intel-6-Westmere-EX .pdf

[164]: Next-Generation AMD Opteron Processor with Direct Connect Architecture – 4P Server Comparison, http://www.amd.com/us-en/assets/content_type/DownloadableAssets/4P_Server_Comparison_PID_41461.pdf

[165]: Supermicro Motherboards, X6QT8, User Manual, https://www.supermicro.com/products/motherboard/Xeon7000/E8501/X6QT8.cfm

[166]: Haas J. & Vogt P., Fully buffered DIMM Technology Moves Enterprise Platforms to the Next Level, Technology Intel Magazine, March 2005, pp. 1-7

[167] Supermicro Motherboards, MNL-1594X6QT8, User Manual, https://www.supermicro.com/manuals/motherboard/C606_602/MNL-1594.pdf

[167]: Birrittella M. S., Intel Omni-Path Architecture Technology Overview, August 2015, http://www.hoti.org/hoti23/slides/rimmer.pdf

[168] Danowitz A. & al. CPU DB: Recording Microprocessor History, Communications of the ACM, Vol. 55 No. 4, Pages 55-63

[169]: Haas, J. & Vogt P., Fully buffered DIMM Technology Moves Enterprise Platforms to the Next Level," Technology Intel Magazine, March 2005, pp. 1-7

[170]: A New Day for the Datacenter, Presentation at AMD Financial Analyst Day, May 16 2017, http://ir.amd.com/phoenix.zhtml?c=74093&p=irol-analystday

[171]: AMD, Intel server market trends, XbitLabs, Aug. 2011, http://www.xbitlabs.com/picture/?src=/images/news/2011-08/amd_intc_server_mkt_trends.png

[172]: Brayton J.,"Nehalem: Talk of the Tock, Intel Next Generation Microprocessor," IDF, April 2008, Shanghai, http://inteldeveloperforum.com.edgesuite.net/shanghai_2008/ti/TCH001/f.htm

[173]: X10QBi Platform User's Manual Revision 1.2, MNL-X10QBi, SuperMicro, 2017 https://www.supermicro.com/products/motherboard/Xeon/C600/X10QBI.cfm

[174]: X11DPT-B User's Manual, SuperMicro, Revision 1.0, Aug. 2017, https://www.supermicro.nl/manuals/motherboard/C620/MNL-2020.pdf

[175]: Intel Secretly Firefighting a Major CPU Bug Affecting Datacenters?, TechPowerUp, Jan. 3, 2018, https://www.techpowerup.com/240174/intel-secretly-firefighting-a-major-cpu-bug-affecting-datacenters#g240174

[176]: Teylor P., Intel shows off Xeon roadmap through 2020, pushes 10nm further back, TechSpot, Aug. 9, 2018, https://www.techspot.com/news/75887-intel-shows-off-xeon-roadmap-through-2020-pushes.html

[177]: Kennedy P., Resetting Scalable Expectations Intel Cascade Lake-SP at Hot Chips 30, ServeTheHome, Aug. 23, 2018, https://www.servethehome.com/resetting-scalable-expectations-intel-cascade-lake-sp-at-hot-chips-30

5. References (22)

[178]: Product Brief: Intel® Optane SSD DC P4800X Series, Intel, 2018, https://www.intel.com/content/www/us/en/solid-state-drives/optane-ssd-dc-p4800x-brief.html

[179]: SSD Cache Intel, 2018, https://www.alza.hu/ssd-cache/intel/18853663-v1294.htm

[180]: Kennedy P., Intel Optane DC Persistent Memory Sampling Today Revenue Delivery 2018, ServeTheHome, May 30, 2018, https://www.servethehome.com/intel-optane-dc-persistent-memory-sampling-today-revenue-delivery-2018/

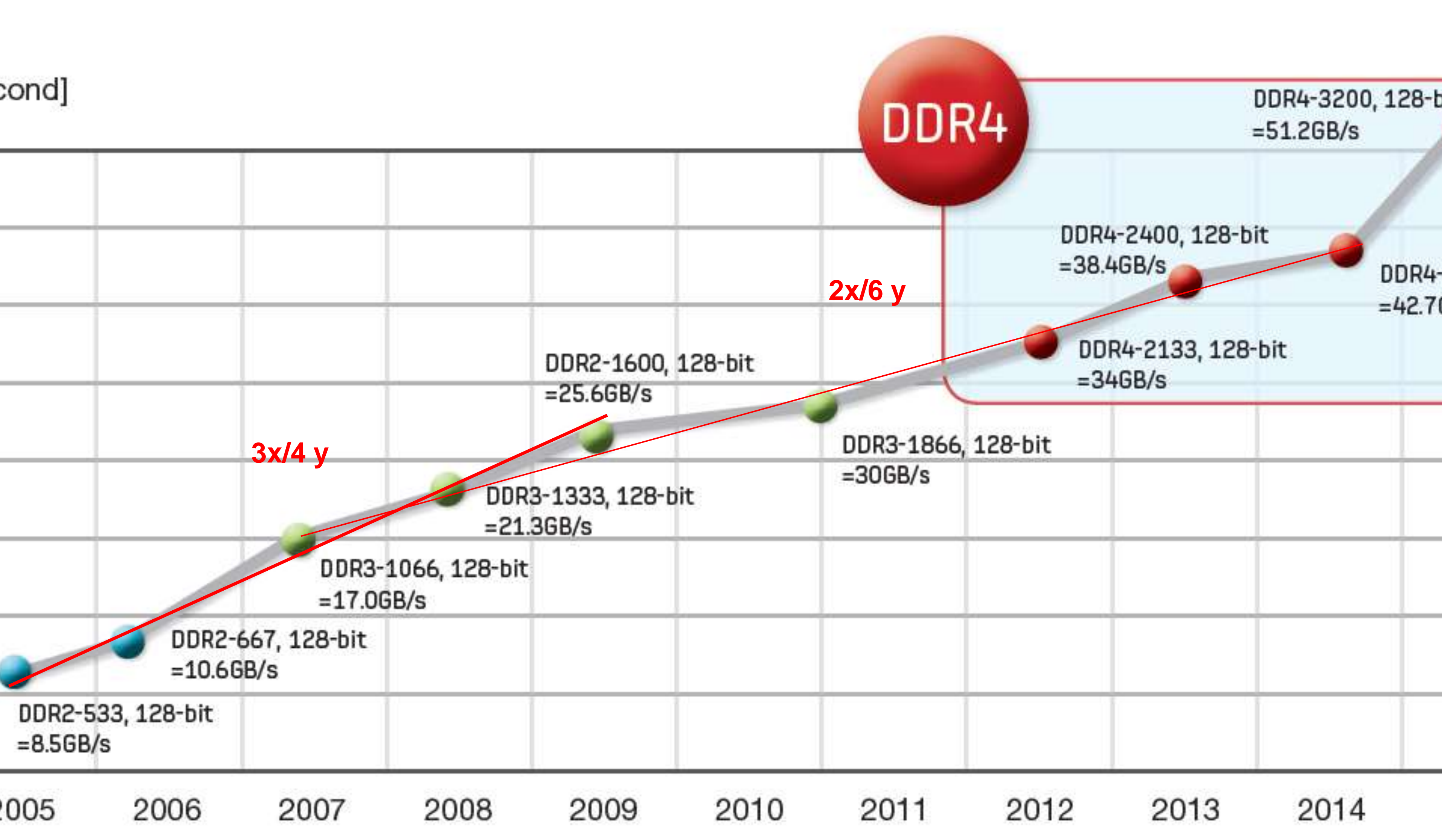
[181]: Reimagining the Data Center Memory and Storage Hierarchy, Intel, Mai 30, 2018, https://newsroom.intel.com/editorials/re-architecting-data-center-memory-storage-hierarchy/

[182]: Cutress I., Intel Launches Optane DIMMs Up To 512GB: Apache Pass Is Here!, AnandTech, May 30, 2018, https://www.anandtech.com/show/12828/intel-launches-optane-dimms-up-to-512gb-apache-pass-is-here

[183]: Tallis B., The Intel Optane Memory (SSD) Preview: 32GB of Kaby Lake Caching, AnandTech, April 24, 2017, https://www.anandtech.com/show/11210/the-intel-optane-memory-ssd-review-32gb-of-kaby-lake-caching/5

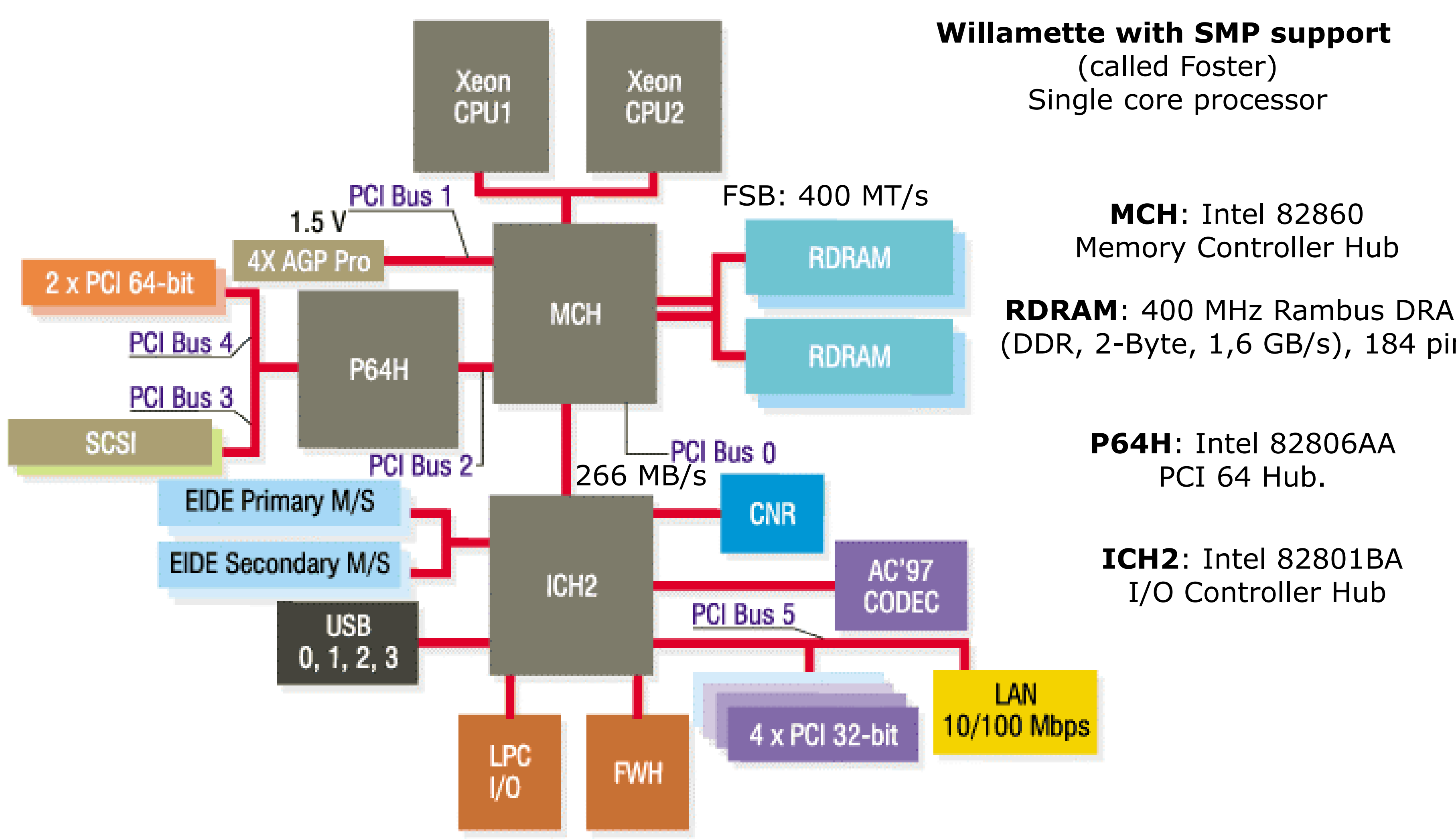
[184]: Tallis B., Intel Optane SSD DC P4800X 750GB Hands-On Review, AnandTech, Nov. 9, 2018, https://www.anandtech.com/show/11930/intel-optane-ssd-dc-p4800x-750gb-handson-review

[185]: Liu Z., Lenovo Lists Intel Cascade Lake Xeon CPU Lineup, TomsHardware, Nov. 10, 2018, https://www.tomshardware.com/news/intel-cascade-lake-xeon-cpu-specs,38053.html



higher performance compared with DDR3L and DDR2

Example block diagram of an early Intel Pentium 4 based DP server platform []



Willamette with SMP support
(called Foster)
Single core processor

MCH: Intel 82860
Memory Controller Hub

RDRAM: 400 MHz Rambus DRAM
(DDR, 2-Byte, 1,6 GB/s), 184 pins

P64H: Intel 82806AA
PCI 64 Hub.

ICH2: Intel 82801BA
I/O Controller Hub

http://ixbtlabs.com/articles/pentium4xeonsmp/

The first PC motherboards with support for RDRAM debuted in 1999. They supported PC-800 RDRAM, which operated at 400 MHz and delivered 1600 MB/s of bandwidth over a 16-bit bus using a 184-pin RIMM form factor. Data is transferred on both the rising and falling edges of the clock signal, a technique known as double data rate. For marketing reasons the physical clock rate was multiplied by two (because of the DDR operation), therefore, the 400MHz Rambus standard was named PC800. This was significantly faster than the previous standard, PC-133 SDRAM, which operated at 133 MHz and delivered 1066 MB/s of bandwidth over a 64-bit bus using a 168-pin DIMM form factor. Moreover, if a mainboard has a dual or quad-channel memory subsystem, all of the memory channels must be upgraded simultaneously. Sixteen-bit modules provide one channel of memory, while 32-bit modules provide two channels. Therefore, a dual channel mainboard accepting 16-bit modules must have RIMMs added or removed in pairs. A dual channel mainboard accepting 32-bit modules can have single RIMMs added or removed as well.

400 MHz DDR 2-Byte resulting in 1600 MB/s 184 pins
PC133 SDRAM 133 MHz 8-Byte resulting in 1066 MB/s 168 pins

http://www.memoryx.com/rdram.html

ARM
(MCs are linked via the
CCN-5xx interconects
to the platform
2012-2014)
16C/32C/48C