Server Sizing 2018

Joe Chang

qdpma.com

Jchang6 at yahoo

About Joe

- SQL Server consultant since 1999
- Query Optimizer execution plan cost formulas (2002)
- True cost structure of SQL plan operations (2003?)
- Database with distribution statistics only, no data 2004
- Decoding statblob/stats_stream
  - writing your own statistics
- Disk IO cost structure
- Tools for system monitoring, execution plan analysis

See http://www.qdpma.com/
ExecStats download: http://www.qdpma.com/ExecStatsZip.html
Blog: http://sqlblog.com/blogs/joe_chang/default.aspx

AWS EC2 Dedicated Hosts

Instance Type	Sockets	Physical Cores
c3	2	20
c4	2	20
c5	2	36
p2	2	36
g2	2	16
g3	2	36
m3	2	20
m5	2	48

	Sockets	Cores
m4	2	24
d2	2	24
r4	2	36
r3	2	20
h1	2	36
i2	2	20
i3	2	36
x1	4	72
x1e	4	72

Mostly 2 sockets (why?)

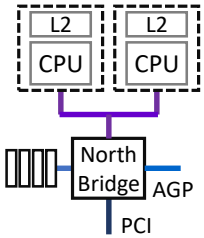
8, 10, 12, 18, 24 cores per socket

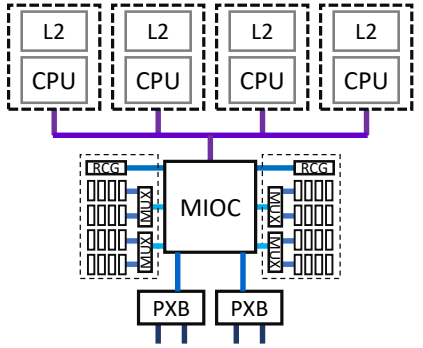
two 4 socket options
18 cores per socket

E5 v3
10 core 2.3, 2.6 & 3.1GHz
18 core 2.3GHz
E5 v4
24 core
Industry Representative

https://aws.amazon.com/ec2/dedicated-hosts/pricing/

The Standard Server Systems





- 2-socket as a standard system for 20 years
  - Original practice was: 2-way for light apps
    - Database: 4-way system with large memory + IO
  - Last 4 years:
    - 2-way is standard system for most uses
    - 4-way for special circumstances
- Why?
  - Originally, there were good reasons
  - Reasons evolved over the years
- But now (recently), its different
  - SSD/Flash storage is more practical than HDD
    - Massive memory overkill no longer necessary
  - Are cores equal between single and 2-socket?

Systems, Processors, Memory

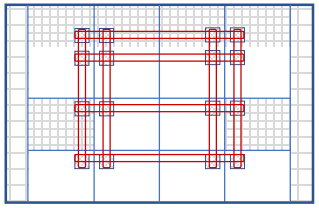
- Modern processors – up to 28 cores Xeon SP
  - Core - very powerful, 2-4GHz, superscalar (8-wide)
    - Hyper-Threading, decode unit alternates between threads
  - 6 memory channels (2 memory controllers)
- Memory
  - DDR4, 2666MT/s (mega-transfers/sec, 1333MHz)
  - Timing CAS (CL), but vague?
    - Fast relative to disk
    - Slow compared to CPU clock
- System – (1), 2 or 4 sockets

HPE 32GB (1x32GB) Dual Rank x4 DDR4-2666 CAS-19-19-19 Registered Smart Memory Kit

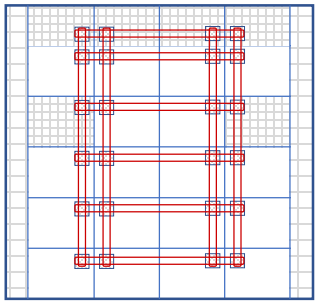
HPE 128GB (1x128GB) Octal Rank x4 DDR4-2666 CAS-22-19-19 3DS Load Reduced Memory Kit

Xeon SP - Skylake

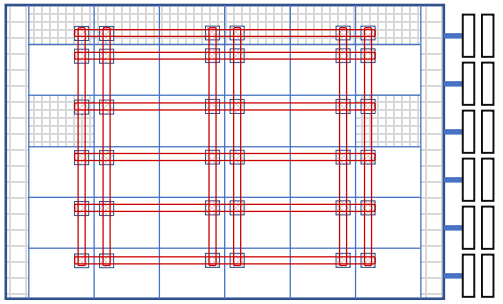
- 13 core count options
  - 4-28 cores
- 3 die layouts
  - LCC, HCC and XCC
  - 10, 18 and 28 cores
- Xeon SP (2017)
  - Now 6 memory channels
    - Still 12 DIMMs
  - 48 PCI-E lanes, 2-3 UPI



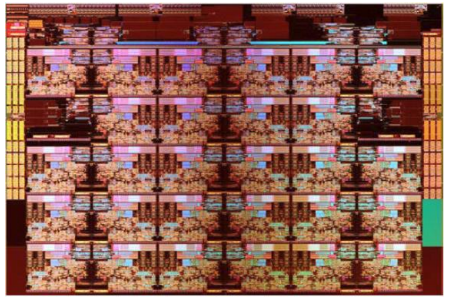
LCC
10 cores



HCC
18 cores



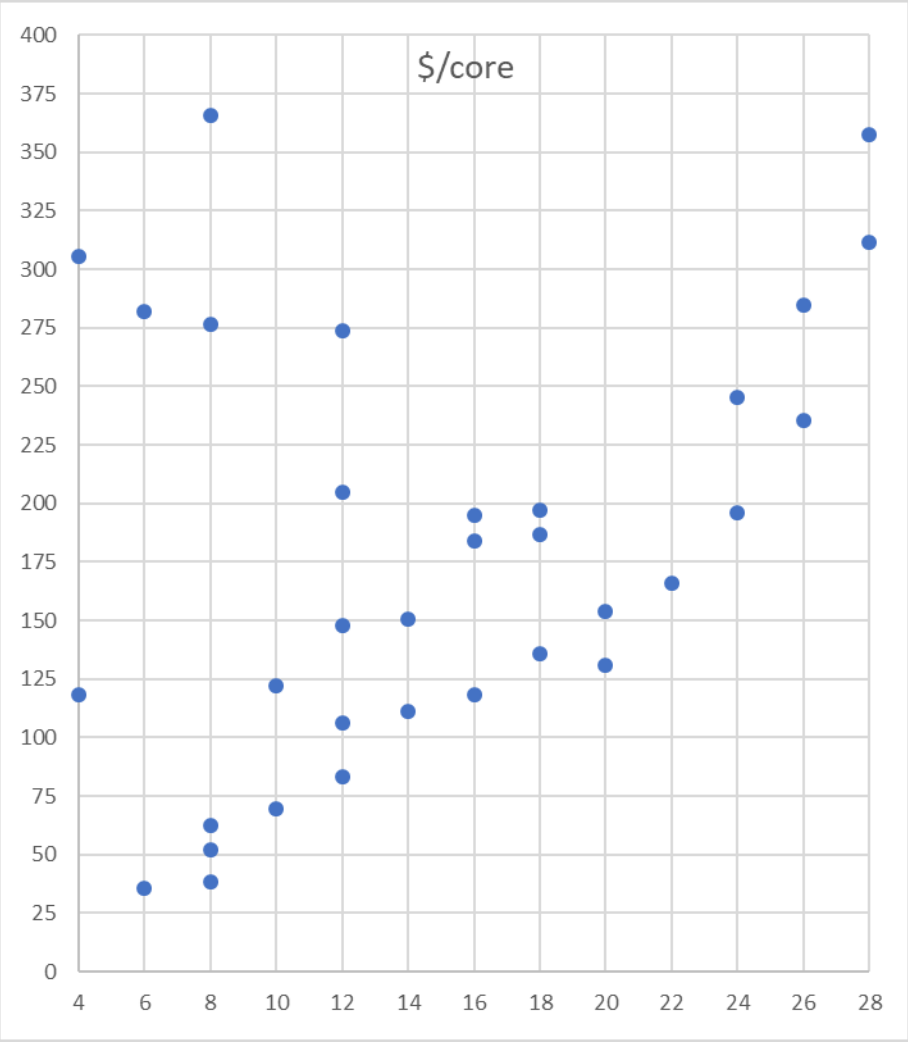
XCC
28 cores



64GB DIMM $1000
128GB DIMM $4000

https://en.wikichip.org/wiki/intel/microarchitectures/skylake_(server)

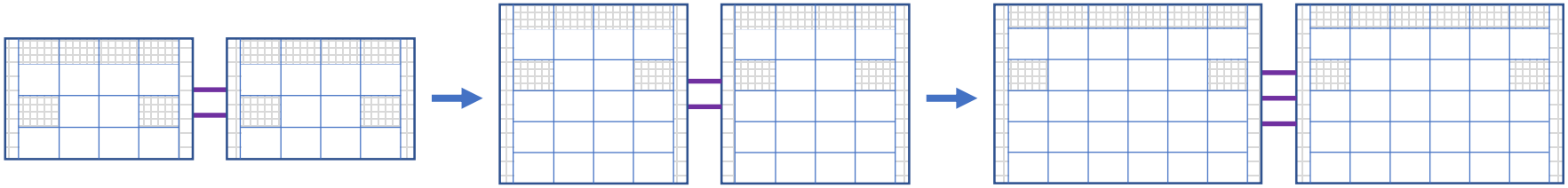
Xeon SP – Price per core vs. cores



8180	28c,	2.5GHz	$10,009
8176	28c,	2.1GHz	$8,719
8168	24c,	2.6GHz	$5,890
8160	24c,	2.1GHz	$4,702
6148	20c,	2.4GHz	$3,072
6138	20c,	2.0GHz	$2,612
6139	18c,	2.3GHz	$2,445
6130	16c,	2.1GHz	$1,894
6132	14c,	2.6GHz	$2,111
5120	14c,	2.2GHz	$1,555
6126	12c,	2.7GHz	$1,776
4116	12c,	2.1GHz	$1,002
5115	10c,	2.4GHz	$1,221
4114	10c,	2.2GHz	$694
41xx	8c	1.7-2.1	$306-501

two 14-core processors less expensive than one 28-core

Standard Systems – first try



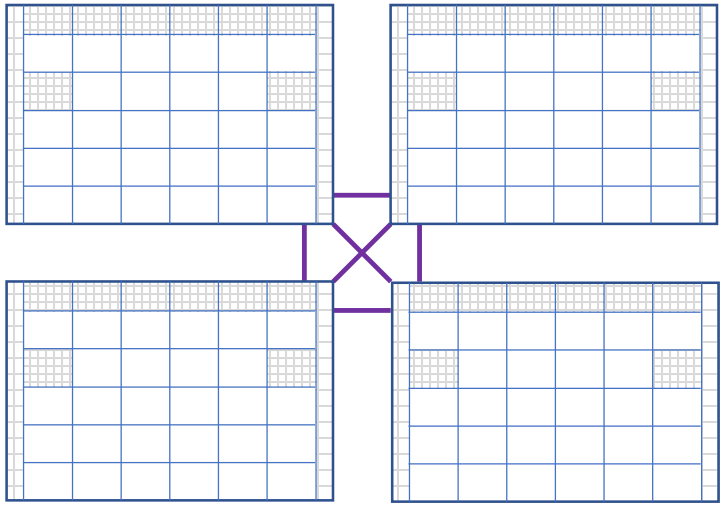
Start with 2-socket LCC, 8 or 10c

– lowest cost per core <$70/c

2-socket HCC (middle)?

– 16 & 18c made from XCC die? $120/c

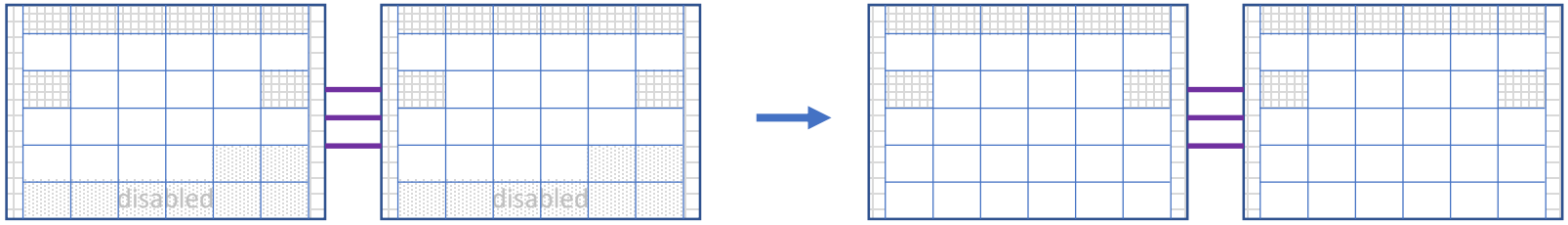
2S XCC – 24c $196/c, 28c $300/c



4S for special requirements?

But, processor cost is only part of system cost,
what happens when system (chassis + motherboard) is included
(assuming fixed memory/core)

Standard Systems – second try



Start with 2-socket, 20c for most economical $/core

Inclusive of system (chassis + motherboard = $1900)

– including rack space cost could change this

4144 - 10c	$69/c + system	= $164/c
6138 - 20c	$130/c + system	= $178/c

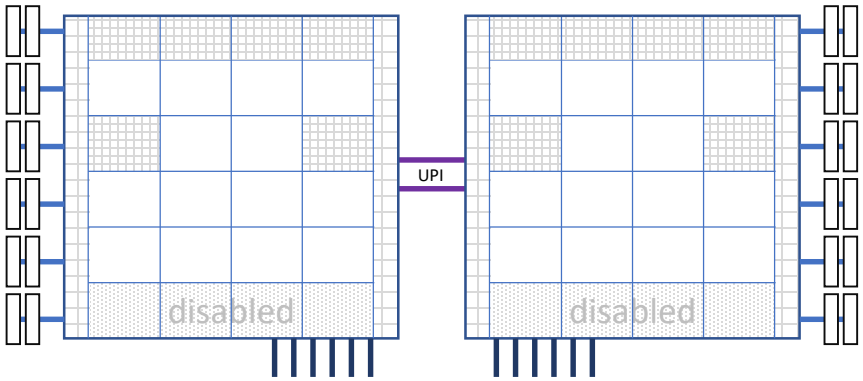
2S 28c for high-end, $345/c

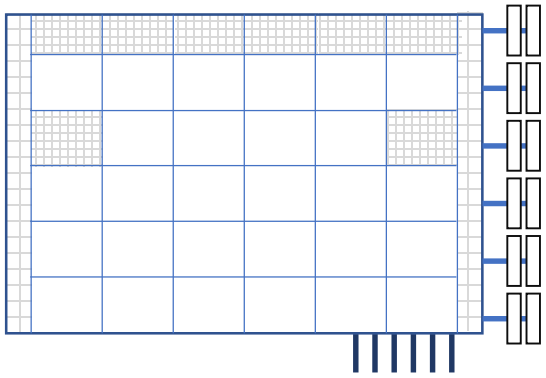
2S high frequency option for some apps (not DB transaction processing)

4-socket if necessary

Disclaimer: do this calculation with your provider

Is 2-Socket a better system?





2 sockets
HCC die (4 cores disabled)
14 cores per socket
28 (active) cores total

6 memory channels per socket
12 DIMMs/socket, 24 total

48 PCI-E lanes/socket, 96 total

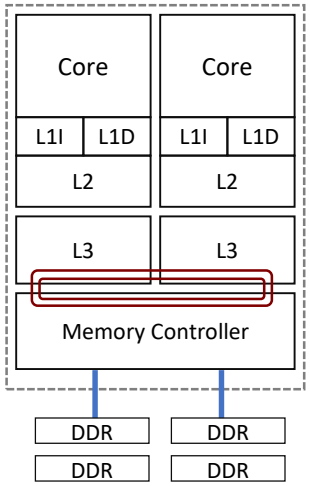
1 socket
XCC die
28 cores total

6 memory channels
12 DIMMs

48 PCI-E lanes

It would seem to be true,
if cores are equal in each system

What Specs are Important?



Desktop quad-core
4GHz+ (0.25ns cycle)
L1 4 cycles

L2 12 cycles

L3 – 42 cycles (10.5ns)
Kaby Lake 38 cycles

51ns at controller?

37.5ns at DDR interface

61.5ns total?

Almost everything in the processor-core has improved dramatically over the past twenty years.

Memory latency has not improved by much, depending

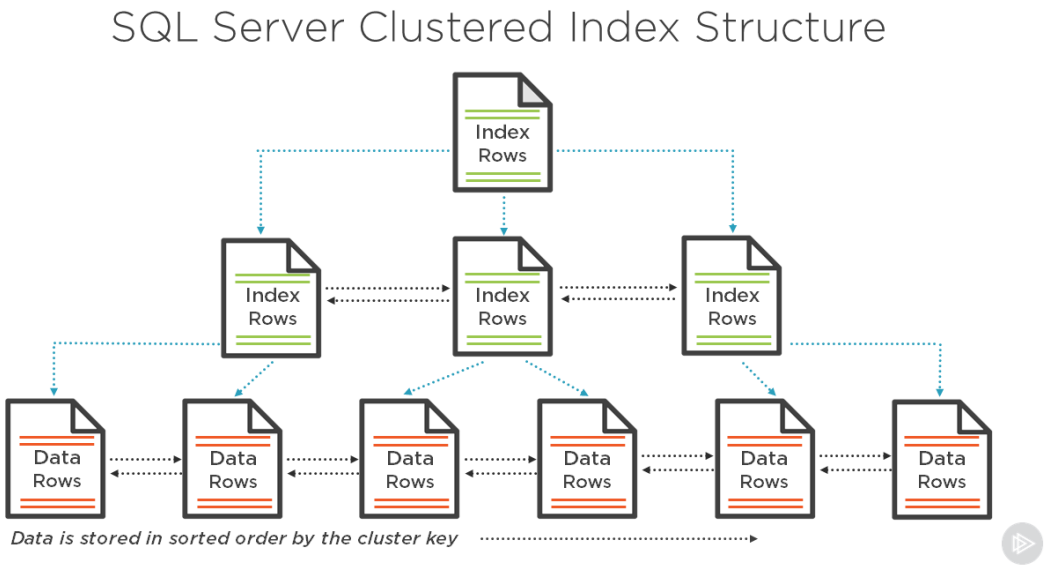
It may now be the most important? because the rest of the processor is so awesome? (core, number of cores, etc.)

Memory bandwidth not particularly important for database transaction processing

Memory capacity was once important, but is now far beyond what is really necessary (with good All-Flash storage)

Memory round-trip is really important

What does DB code do?



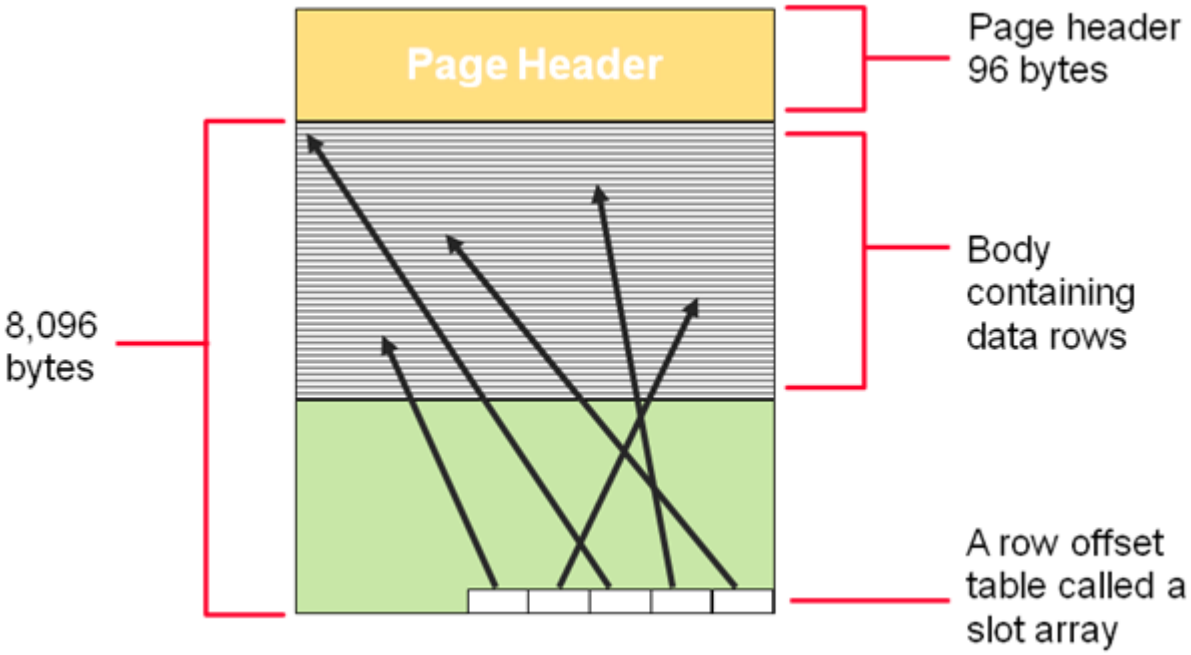
Navigate system tables, IAM
Find locator for index root
Read index root page
Find locator for next level
Read index intermediate level
Find locator for leaf level
Read leaf level

In other words
Access memory address
Which points to next address

IOT: Pointer chasing

http://buildingbettersoftware.blogspot.com/2016/07/what-about-indexing-sql-server-tables.html

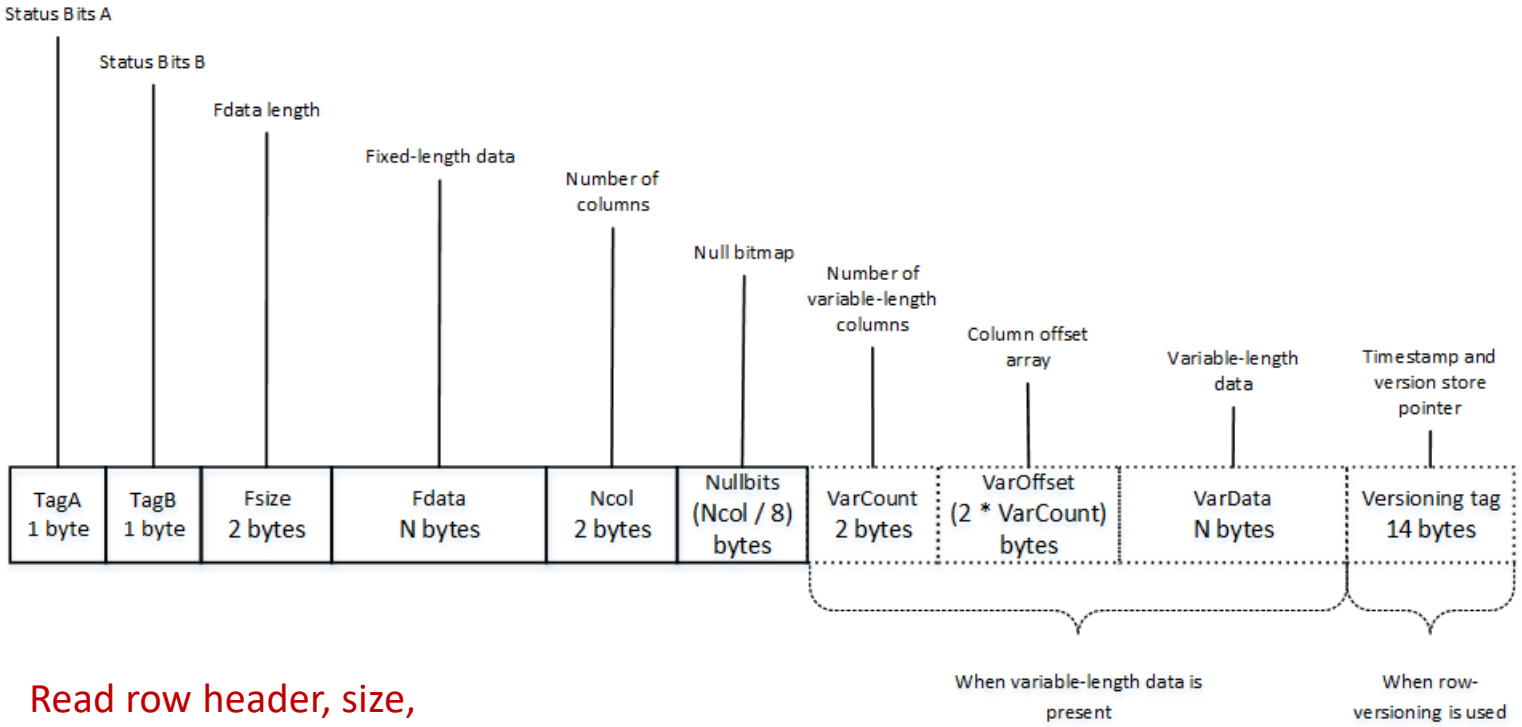
Finding a row within a page



Read row offset at end of page to determine individual row locations then read row

http://forsharingknowledge.blogspot.com

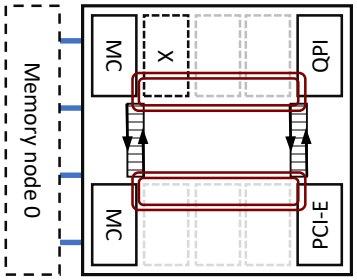
Accessing columns within a row



Read row header, size,
sys.columns for column info
Fixed length (not null) data locations are known
Get Null bitmap, column offset array
to locate individual variable length columns

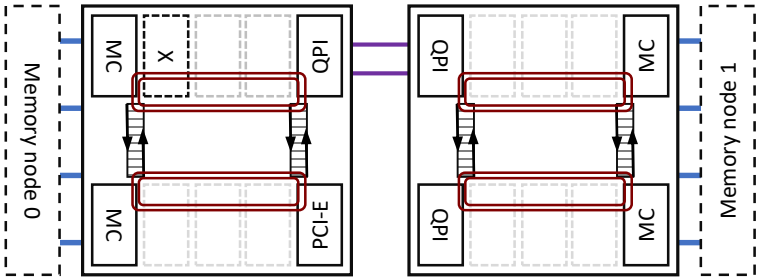
http://Aboutsqlserver.com

Memory Latency – 1S & 2 Socket



Single socket
All memory local
On L2 miss
Local L3: = 17-18ns

Mem: L3 + 58ns = 76ns
(Xeon E5-2630 v4, 10-core)



Local node, 50% Remote node, 50%

Local node L3: 17-18s,
Remote node L3: 30-100ns?

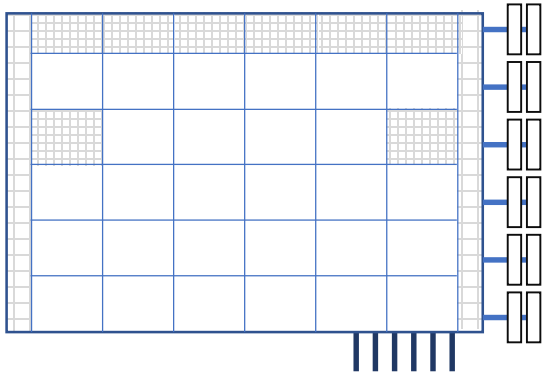
Local memory: L3 (18ns) + 75? = 93ns
Remote memory (50-60ns longer)? = 148ns?

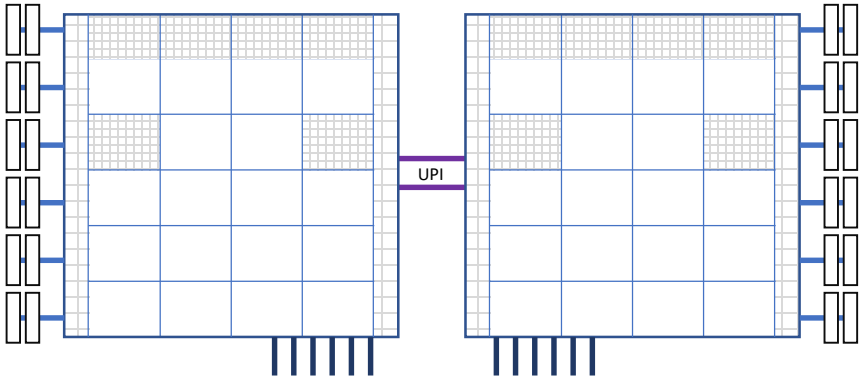
Single socket based on
Xeon E5-2630v4, 10-core
(single ring)

2-socket
Xeon E5-2699v4, 24-core

Measurements on matching systems needed

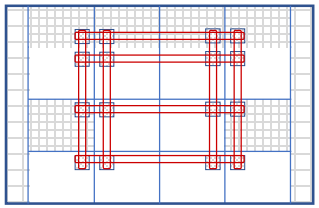
Xeon SP





All memory local

L3 cache 18-19ns?
DDR 50ns?
Tot. Mem 70ns?

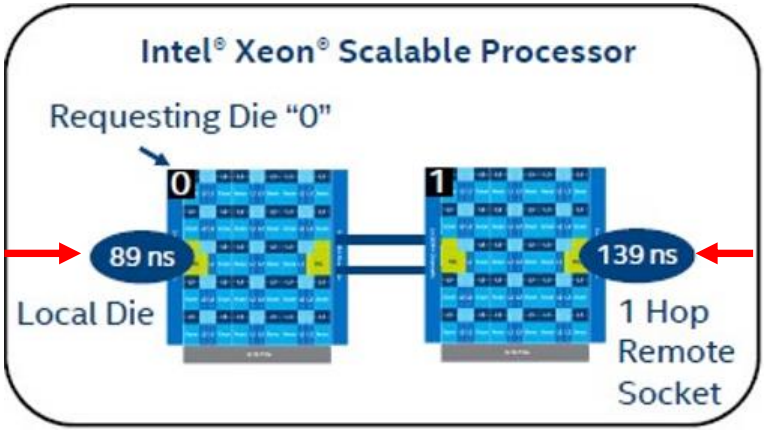


http://7-cpu.com/cpu/Skylake_X.html
DDR4 3400 16-18-18-36

Probably 75ns for ECC RDIMM at 2666-19

Memory access
50/50 local remote

Local node 89ns
Remote node 139n



https://www.nextplatform.com/2017/11/27/intel-stacks-xeons-amd-epyc-systems/

Hypothetical Transaction

2.5GHz processor core, 0.4ns per cycle
Suppose a transaction consists of 3M operations (that complete in a single CPU-cycle)
If memory access is single cycle (disregard L2 and L3 at 12 and 50 cycles)
Then 3M ops completes in 1.2ms, performance is 833 transactions per/sec

Real memory – 76ns (1-socket), 190 cycles on 2.5GHz core
Suppose 2.456% of operations involve waiting for a round-trip memory access
2.925M ops complete in 1 cycle, 75K ops incur 190 cycle wait
Total cycles (work + wait) = 16.9M, Performance 147.7 tps

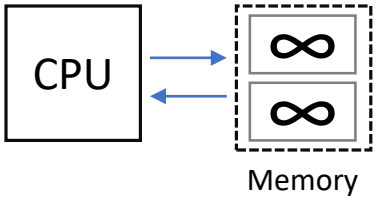
2-Socket – 50/50 local remote split, 93ns local, 148ns remote
120.5ns average memory, 301 cycles
Performance per core 99.5 tps, 32.5% lower, 1.35X system level gain

Skylake parameters
1S 81.5ns (204 cycle), 2S 89/139 local remote, 114ns average (285)

1S, 75K ops incur 204 cycle wait, Total cycles (work + wait) = 17.94M, Performance 139.4 tps

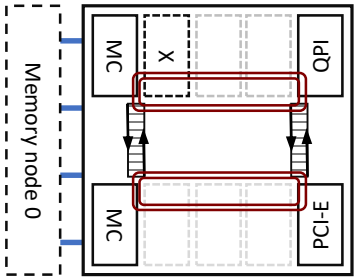
2-Socket – 50/50 local remote split, 89ns local, 139ns remote, 114ns average memory, 285 cycles
Performance per core 104.5 tps, 25% lower, 1.5X system level gain

Hypothetical Tx – 2.5% memory access



Core 2.5GHz, 0.4ns/clock
Transaction: 3M operations
memory 1 cycle

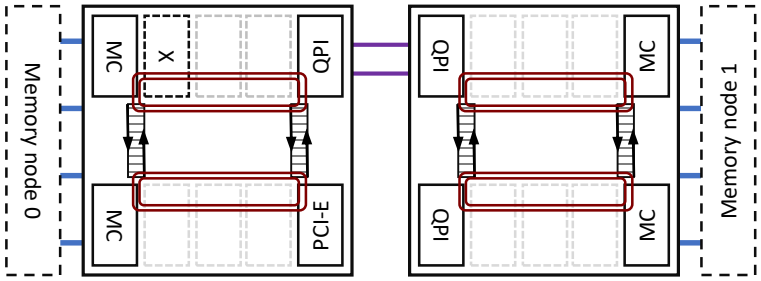
Perf: 833 tx/sec



memory 76ns or 190 clocks
2.5% of ops – mem access, 0.025 * 3M = 75K

Single cycle op 2.925M
75K (mem) x 190 = 14.25M
Tot cycles 17.175M

Perf/thread: 145.56 tx/sec

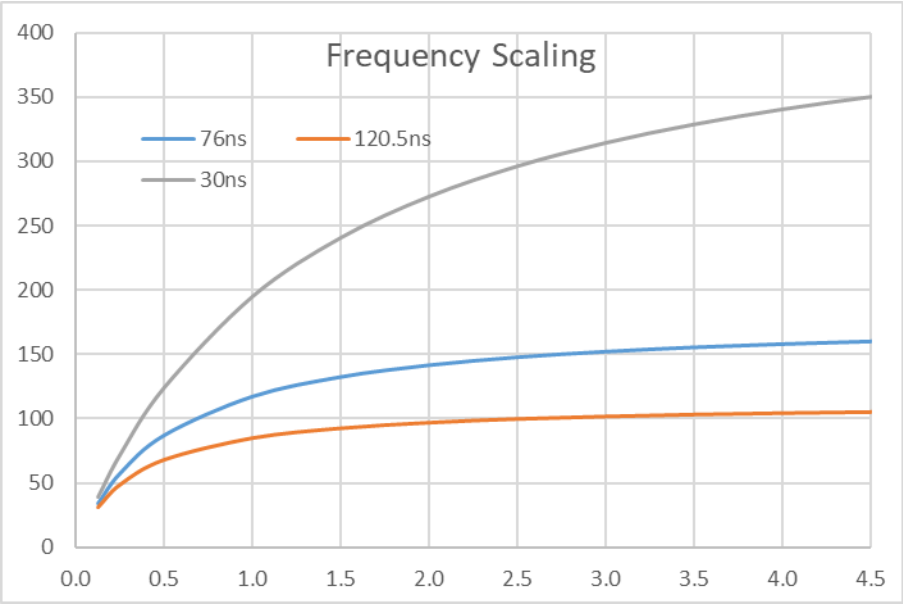


Avg. mem (93 + 148)/2 = 120.5ns = 301 clocks

Single cycle op 2.925M
75K (mem) x 301 = 22.575M
Tot cycles 25.500M

Perf/thread: 97.97 tx/sec

Frequency Scaling



Any facts to support this?

Test system
2-socket Xeon E5-2680 (v1)
8-cores, 2.7GHz

BIOS/UEFI update
System set to power save
135MHz (20X reduction)

CPU-cycles increase by 3X

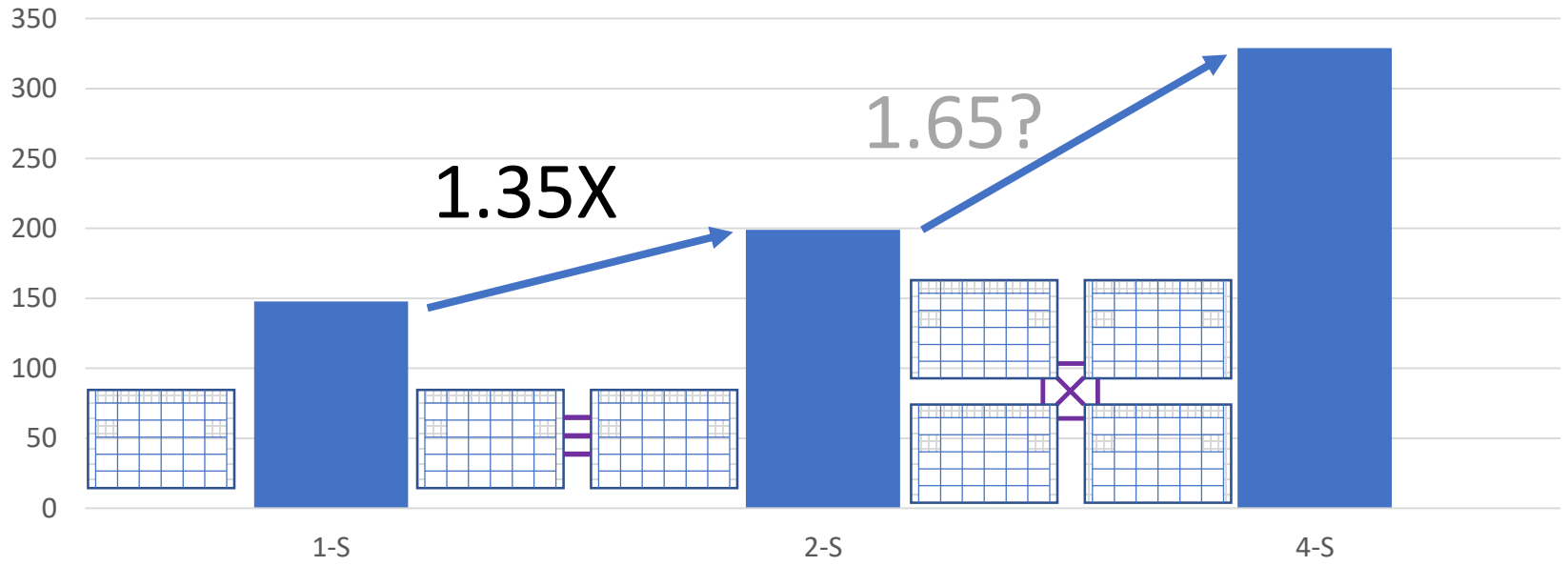
Simple model,
2.5% operations wait for round-trip memory access,
76 ns – 1S,
120.5ns – 2S 50/50 local/remote, 93/148ns
(30ns SRAM as main memory)

Hyper-Threading / SMT

- Each physical core pretends to be 2 logical processors (IBM POWER 8/9 – 4 or 8 way SMT)
- Intel implementation
  - Decode unit switches between threads

- This model can achieve linear scaling only if
  - there are many dead cycles for a thread/LP
  - There are many (85-90%) in transaction processing

Scaling versus Sockets



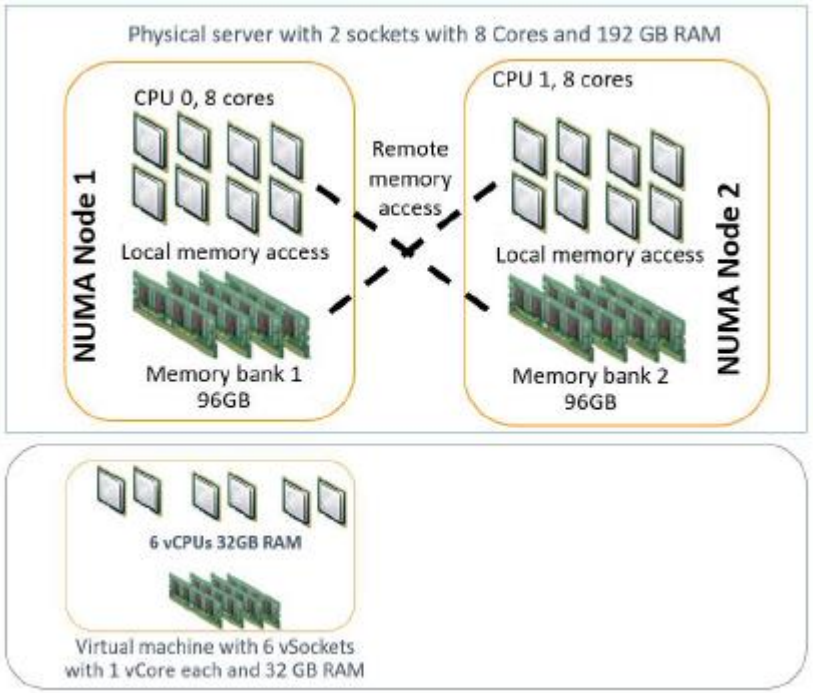
Hypothetical transaction, 3M real CPU-operations (not no-µops), 2.5% incur round-trip memory,

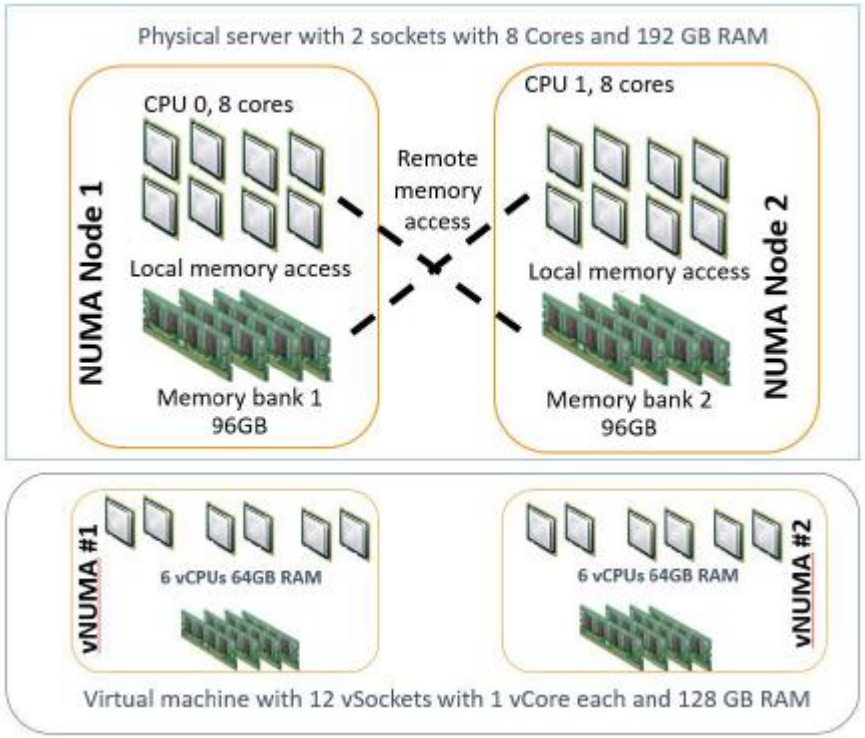
- 1-socket, all memory local (75ns)
- 2-socket, 50/50 local/remote mem (93/148ns, avg. 120.5)
- 4-socket, 25/75 local/remote w/SMB (108/163)

1-S 3GHz, 152tx/s,
7.4% increase over 2GHz

TPC-E 2S – 4S 1.72X scaling

SQL Server on VMware





VMware uses their own terminology, read their document carefully, verify it is what you think it means?

Architecting Microsoft SQL Server on VMware vSphere

Hard Evidence

- Frequency
  - 135MHz – 2700MHz, 20X (2S E5 v1)
  - 3X performance
- 2-Socket to 4-Socket
  - SQL statement CPU about 30% higher
- Hyper-Threading – almost linear scaling
- HP paper, NUMA tuning 30-40% ⟵
  - DB & App must be architected together for NUMA

1S versus 2S local memory ~ 17ns (22%) difference
1S versus 2S average latency (50/50) – 58%

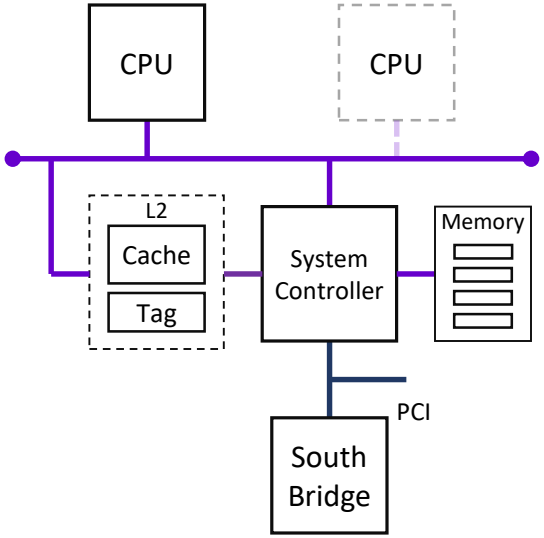
Server Strategy Today: 1-Socket

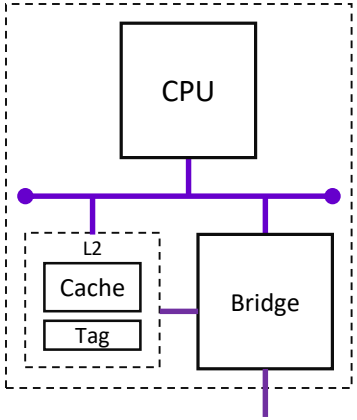
- Sufficient cores for most transaction workloads
- Sufficient memory capacity 12 x 64GB = 768GB
  - With SSD/flash storage,
  - no need for massive overkill on memory
- Better single thread performance than 2+ sockets
- Few adverse NUMA effects
  - High volume Inserts into single table with identity key
- SQL Server licensing costs dominates (EE) even SE?

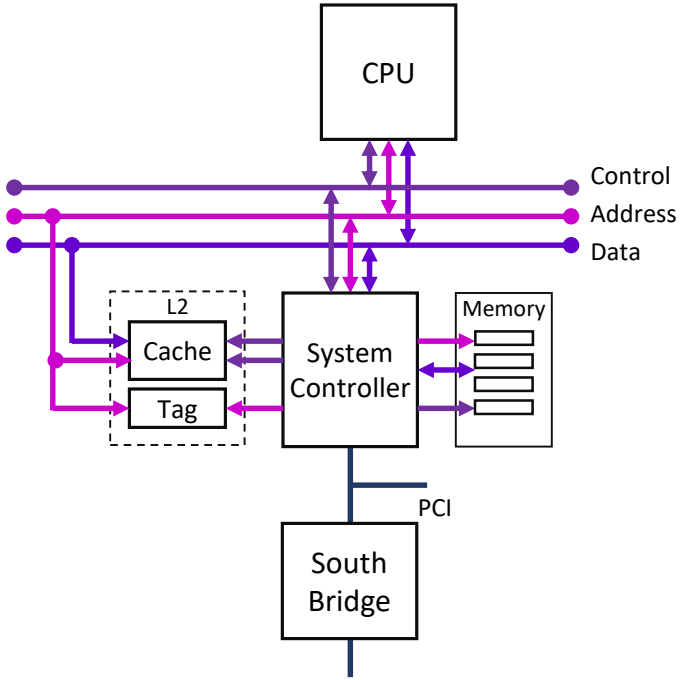
Oracle ROWID – fewer memory round-trips?

Historical

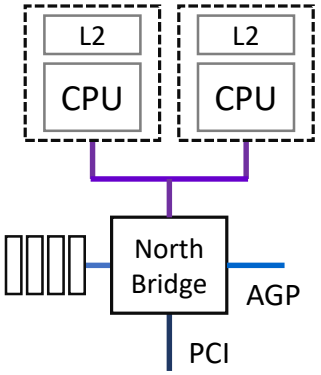
Intel Pentium



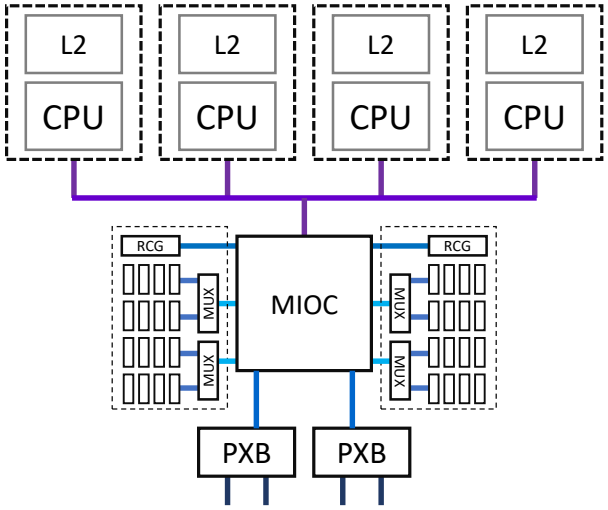




1997-8 Pentium II era



2-way system
– same Pentium II/III processor as desktop,
– desktop chipset (ECC enabled, 4 DIMMs, 1 PCI + AGP)
slightly higher cost structure than single processor system.
Wide adoption of dual-processor by default,
Good for app servers
Inadequate memory and IO for databases

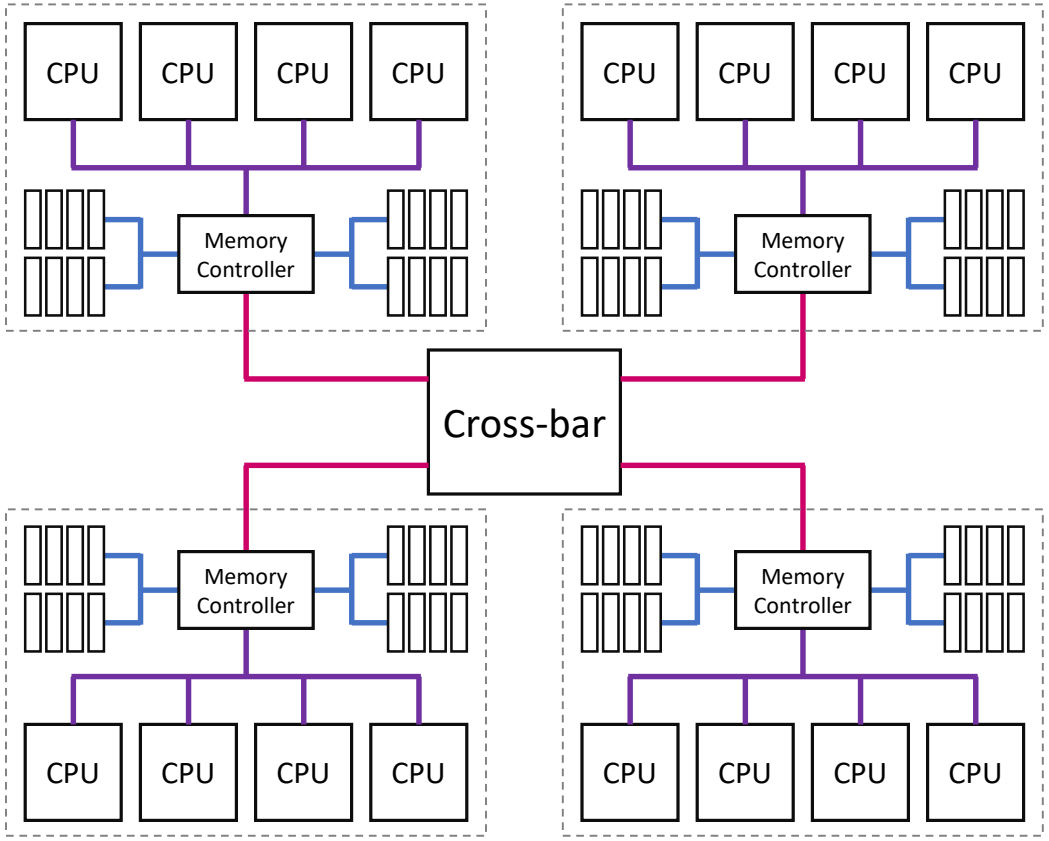


4-way system
– Pentium II/III Xeon processors
– Better cache, less traffic on memory bus (FSB)
Chipset – large memory capacity + multiple PCI
Standard system for databases

But why were these two the standard systems?

Pentium Pro system slightly more complicated, documents difficult to find

Non-Uniform Memory Access **NUMA**



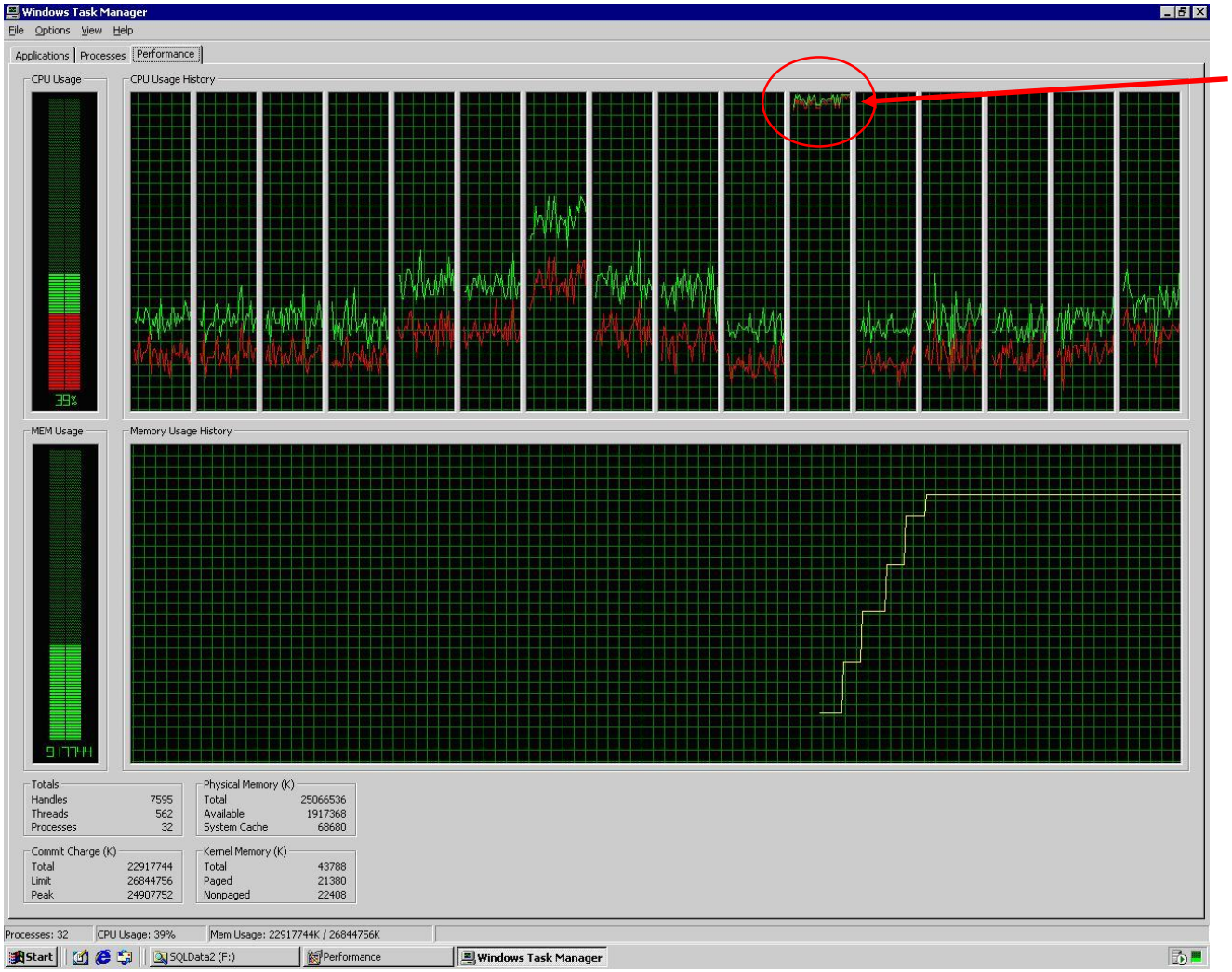
4-way is great, but …
I want Big Iron!

Memory access non-uniform
could be local to a processor node, or on a remote node

What's the big deal?
Software not designed for non-uniform memory
Stupid things can happen

Scaling is possible
However, certain operations exhibit negative scaling, sometimes severely so.
Important: Identify bad operations, then code around it.
Unfortunately, this very little detail was not explained upfront?

SQL Server 2000 – 16-way - SAP



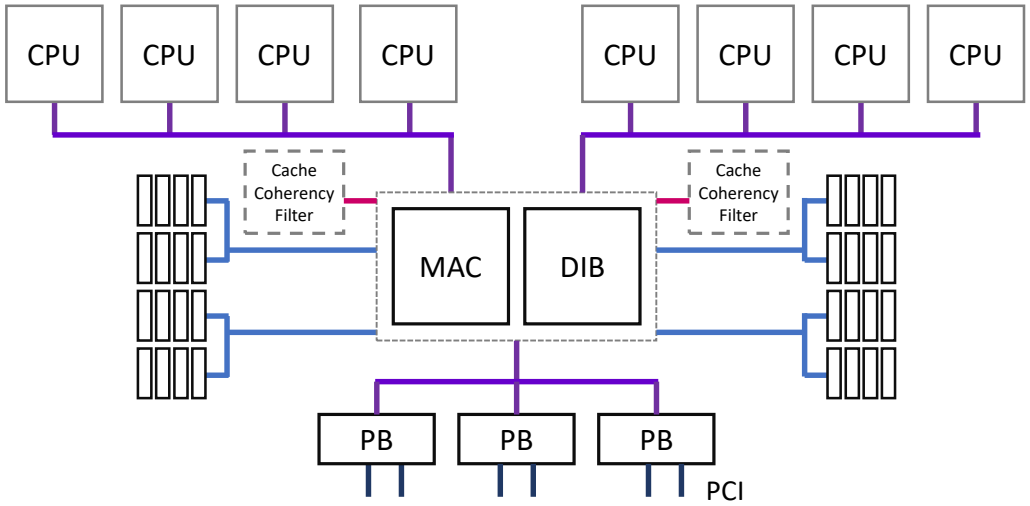
Long ago
Network interrupts handled by single processor (affinity)

On 4-way, 27K RPC/sec

On 16-way – 12K/sec?

Since then
Intel Extended Message Signaled Interrupts (MSI-X) and many other techniques

ProFusion - 1999



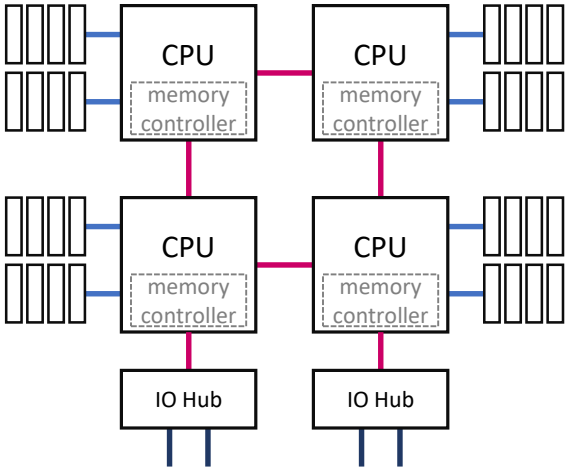
Point product – Corollary
Acquired by Intel
8-way chipset for Pentium III

Was not continued for Xeon (desktop Pentium 4)

Compaq/HP did their own version for Xeon MP

SQL Server did not seem to have adverse issues on the 8-way ProFusion
MS Exchange 5.5 did exhibit unusual effects

AMD Opteron - 2004



Memory access latency

Local node	~60-70ns
Remote node	~100ns?

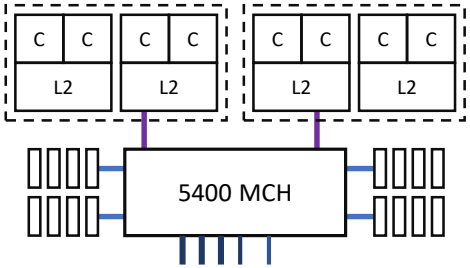
Memory controller integrated into CPU die
Hyper-Transport – point to point protocol between processors and IO Hub

Multi-processor system inherently has non-uniform memory access

but absolute memory latency is low
Remote node latency is comparable to memory on north bridge

AMD emphasized memory bandwidth scaling with nodes.
Excellent performance in applications sensitive to memory latency (OLTP).
(Bad) NUMA characteristics were muted?
was not important, needed 2P for app, 4P for database

2006 – Intel Core 2 era



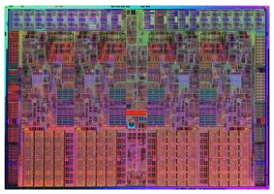
Between 1998 and 2006
2-way gradually acquired a proper server chipset, with adequate memory (16 DIMM slots) and IO (28 PCI-E lanes + DMI)

Quad-core processor, 8 cores total - good enough for medium database workloads

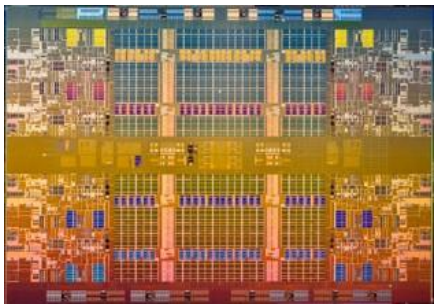
Intel did not have best chipset for the 4-way system in this time
The Opteron 4-way with 32 DIMM slots was a good choice

Intel did not have good 2-way server/workstation chipsets in the Pentium III-4 period. ServerWorks fill the market gap until Intel finally got around to it around 2005?

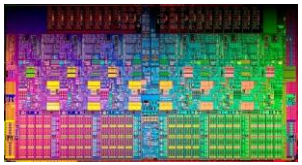
Recent History 2009-11



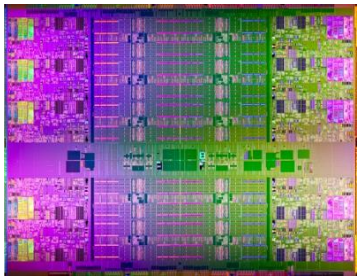
Nehalem
4 cores
(2009)



Nehalem-EX
8 cores (2010)



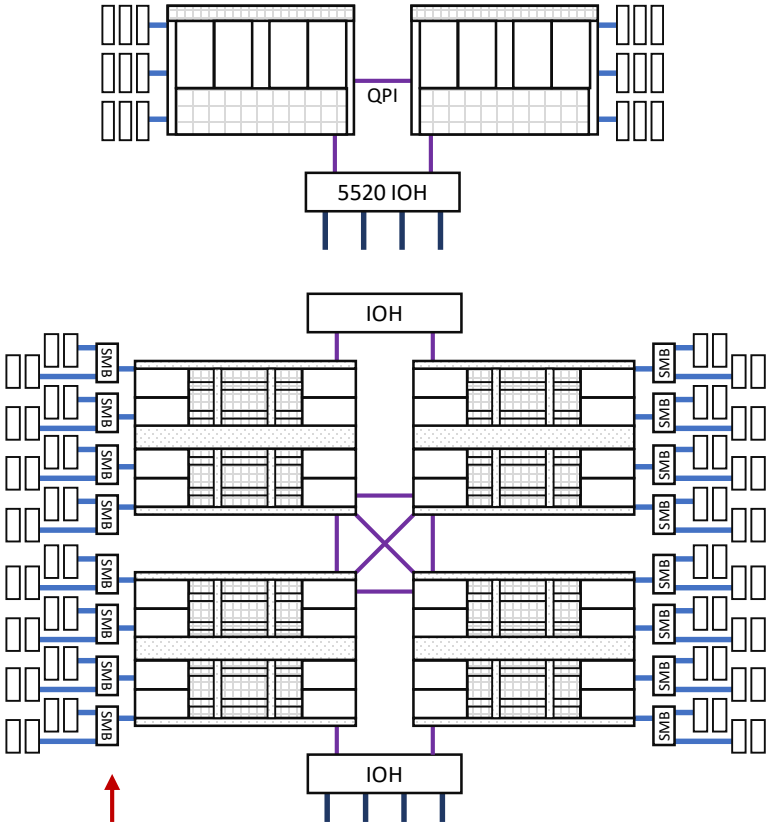
Westmere
6 cores
(2010)



Westmere-EX
10 cores
(2011)

- Little processor for 2-way
- Big processor for 4-way

Nehalem & Westmere (2009-11)

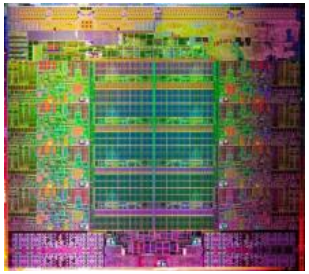


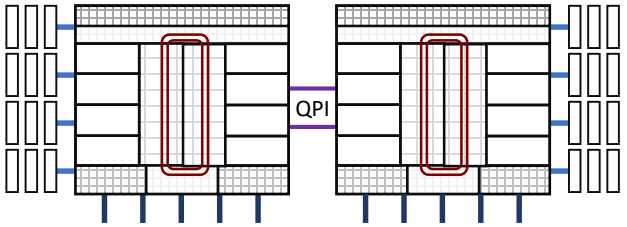
EX model connects to memory via Scalable memory buffer (SMB), (doubles) capacity
adds memory latency and cost

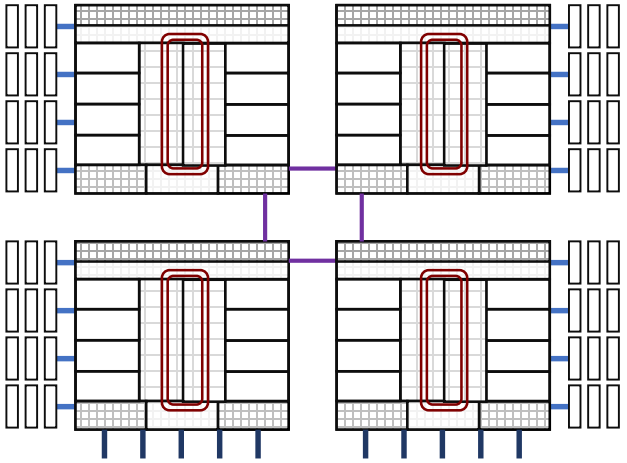
- Little proc for 2-way
  - Processor connects directly to memory
- Big processor for 4-way
  - Scalable Memory Buffer
  - Doubles capacity
  - Adds cost to platform
  - Higher memory latency

http://www.ece.umd.edu/~blj/papers/isca2012.pdf

Intel Sandy Bridge (2012)

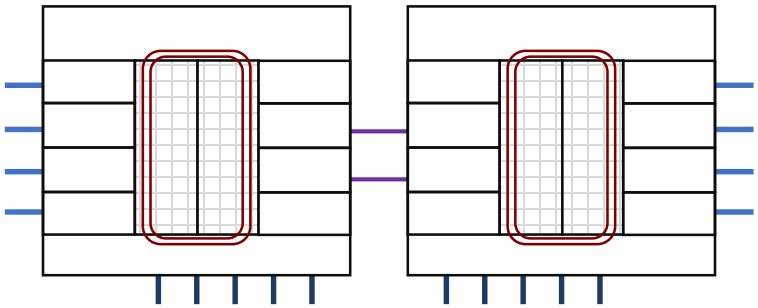






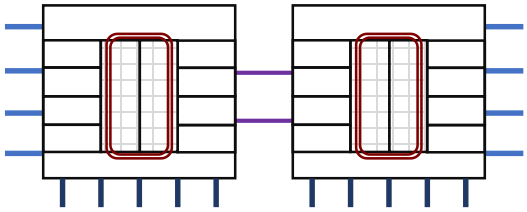
- Sandy Bridge (Xeon E5)
  - for 2 and 4 sockets
  - Xeon E5 2600 and 4600 series
- 4 memory channels
  - 2 controllers
  - 12 DIMM slots
- 40 PCI-E lanes

2012: 2-way



Each socket:
4 memory channels
40 PCI-E lanes + DMI

- Intel Sandy Bridge (EP)
  - 8 cores
  - 4 memory channels (2 controllers)
    - 12 DIMM slots
  - 40 PCI-E lanes + DMI (4 lanes)
- 2012
  - 8 cores is probably good enough with excellent tuning
    - But 2 x 8 cores is better?
  - 8-12 DIMM slots
    - 32GB DIMM $4K in 2012, $1400 in 2013
    - 16GB DIMM $1K in 2012?

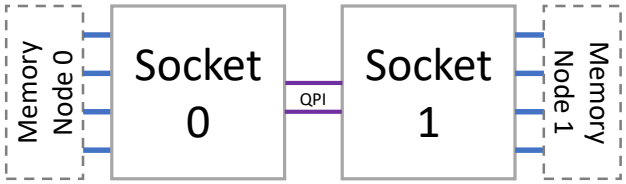


Each socket:
4 memory channels
40 PCI-E lanes + DMI

Single socket 192GB practical memory capacity (16GB DIMMs)
Good but maybe insufficient to reduce IO to noise level
SSD still too expensive for broad use ($10-20/GB?)

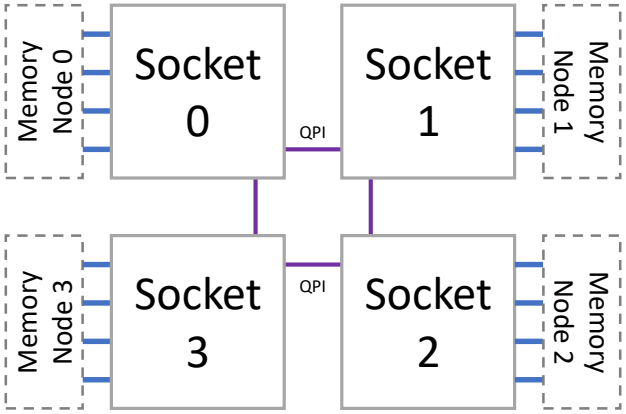
2013-16 Xeon E5/E7 v2, 3 & 4

2 Sockets, Xeon E5 v4



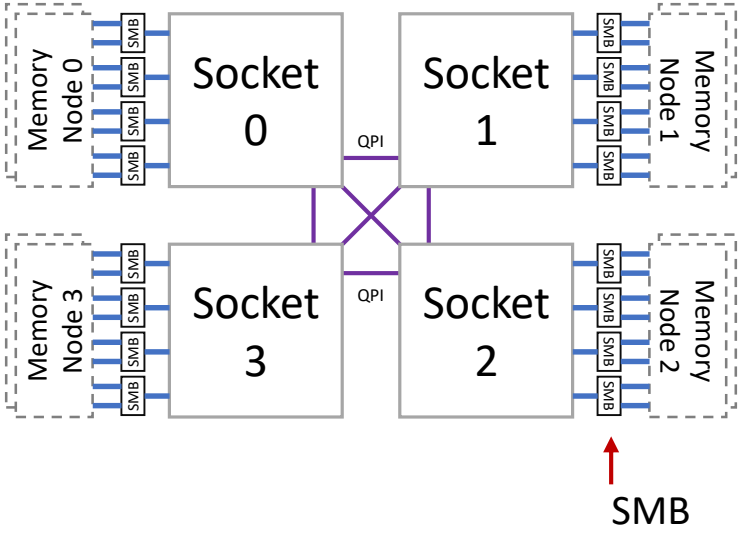
4 Sockets

Xeon E5 v4600



or

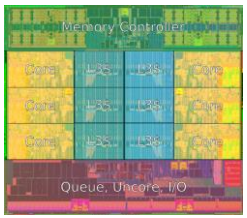
Xeon E7

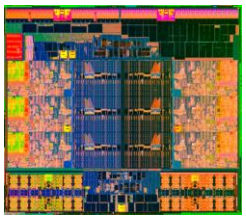
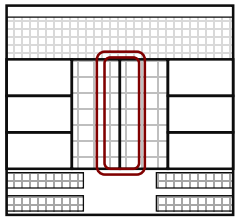
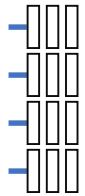
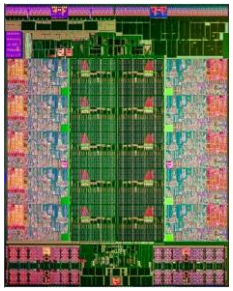
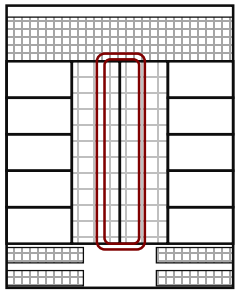
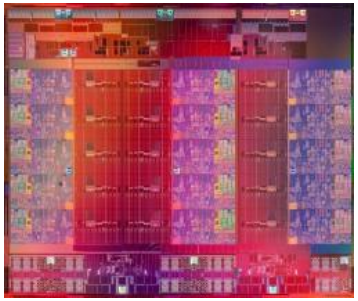
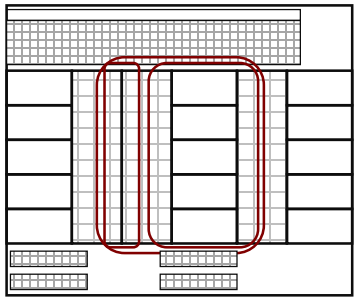


~2015 SSD more practical than HDD for general use?

Ivy Bridge 2013/14

- 3 die layout options
  - LCC, MCC and HCC
  - 6, 10 and 15 cores
- Xeon E5 v2 (2013)
  - 40 PCI-E lanes, 2 QPI
  - Connects directly to memory
  - Limited to 12 cores except AWS etc.
- Xeon E7 v2 (2014)
  - 32 PCI-E lanes, 3 QPI
  - Connects to SMB for memory expansion



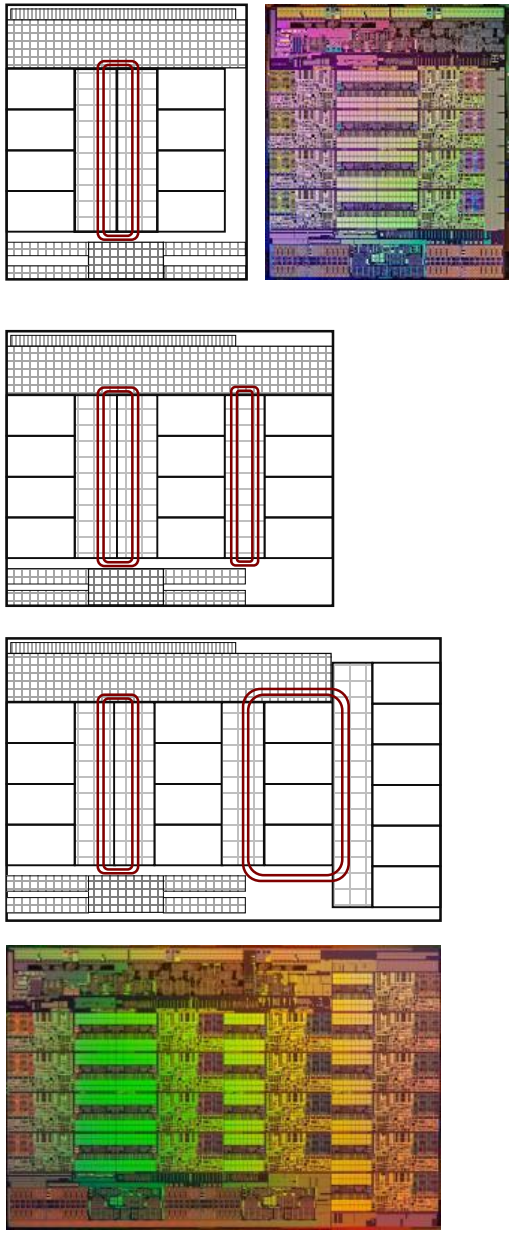








32GB DIMM, $1400 in 2013, 12 x 32GB = 384GB
SSDs somewhat expensive, becoming less expensive

Haswell 2014/15

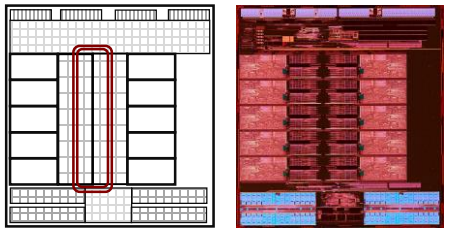
- 3 die layout options
  - LCC, MCC and HCC
  - 8, 12 and 18 cores
- Xeon E5 v3 (2014)
  - 40 PCI-E lanes, 2 QPI
  - Connects directly to memory
  - Limited to 15 cores except AWS etc.
- Xeon E7 v3 (2015)
  - 32 PCI-E lanes, 3 QPI
  - Connects to SMB
- SSDs for general use?

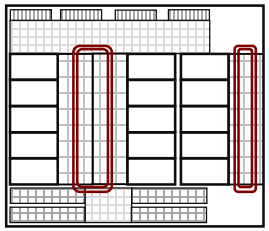
10 TB SSD at $4-6/GB or 100-200 HDDs @ $600 each

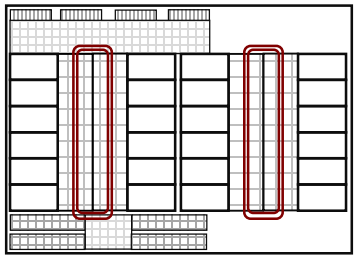


Broadwell 2016

- 3 die layout options
  - LCC, MCC and HCC
  - 10, 15 and 24 cores
- Xeon E5 v4
  - 40 PCI-E lanes, 2 QPI
  - Connects directly to memory
  - Limited to 22 cores except AWS etc.
- Xeon E7 v4
  - 32 PCI-E lanes, 3 QPI
  - Connects to SMB
- SSD for general use





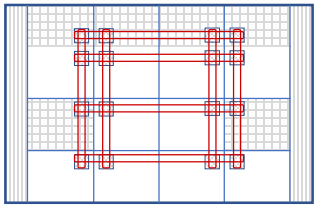


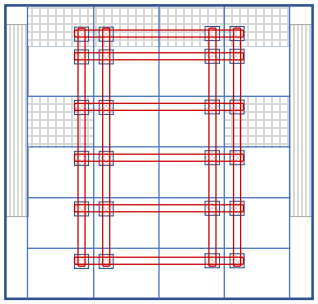
64GB $4000?, 2017 $1000
32GB DIMM $900?

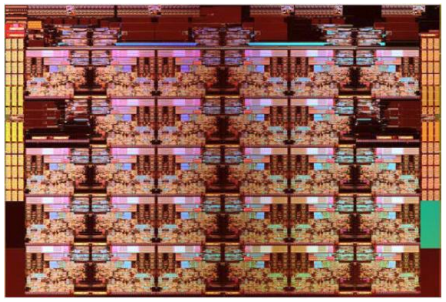
Broadwell 456, 306, 246mm2
25.2 x 18.1mm 17.90 x 12.85 – 71%
18.9 x 16.2mm 13.32 x 11.50
15.2 x 16.2 10.8 x 11.5

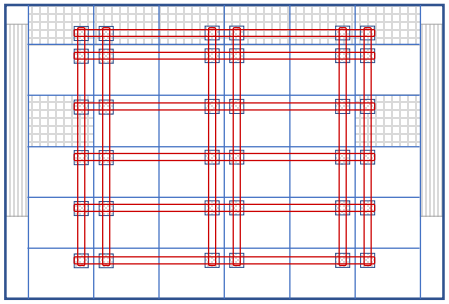
Skylake 2017

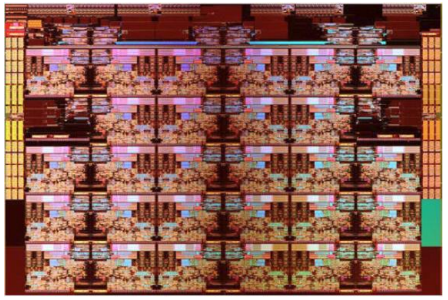
- 3 die layout options
  - LCC, HCC and XCC
  - 10, 18 and 28 cores
- Xeon SP (2017)
  - Now 6 memory channels
    - Still 12 DIMMs
  - 48 PCI-E lanes, 2-3 UPI







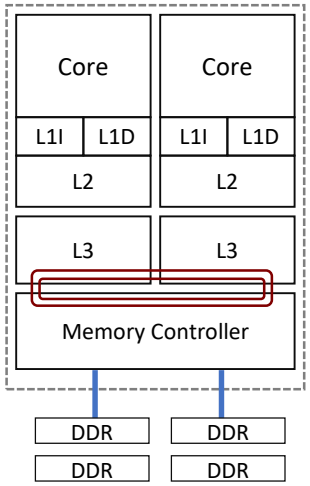




64GB DIMM $1000
128GB DIMM $4000

Memory Latency

7-zip LZMA Benchmark



Desktop quad-core
4GHz+ (0.25ns cycle)
L1 4 cycles

L2 12 cycles

L3 – 42 cycles (10.5ns)
Kaby Lake 38 cycles

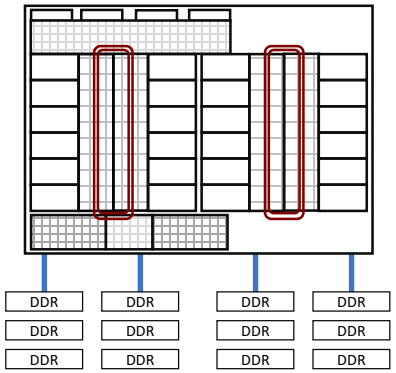
51ns at controller?

37.5ns at DDR interface

61.5ns total?

DDR4-2400 MT/s, CL15
CMD: 1200MHz, T_{RP}+ T_{RCD}+ T_{CAS} = 37.5ns

Ballistix Elite 3000, 15-16-16

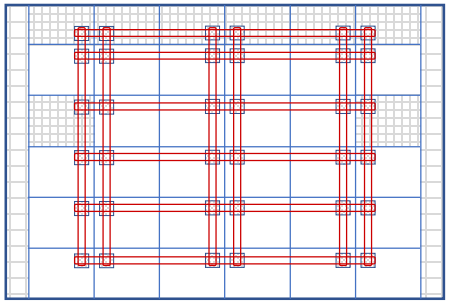


Broadwell E5-2699 v4 (2S)
2.4GHz – 0.417ns cycle
L1 4 cycles
L2 12 cycles

L3 – 60-69 cycles (in turbo?)
18 ns (at base? 43 cycles?)

RAM: 65 cycles (18ns) + 75ns (2S)
18ns + 62ns (1S)

DDR4-2133 MT/s, CL15 (RDIMM, ECC)
CMD: 1066MHz, T_{RP}+ T_{RCD}+ T_{CAS} = 42.2ns



Skylake Xeon SP (1S, 8c)
2.5GHz – 0.4ns cycle
L1 4 cycles
L2 14 cycles

L3 – 68 cycles 3.6GHz
19.5 ns?

RAM: L3 + 50ns?

DDR4-2600 MT/s, CL19 (LRDIMM, ECC)
CMD: 1300MHz, T_{RP}+ T_{RCD}+ T_{CAS} = 43.8ns
CL17 1200MHz, T_{RP}+ T_{RCD}+ T_{CAS} = 42.5ns

http://7-cpu.com/
http://7-cpu.com/cpu/Skylake.html
http://7-cpu.com/cpu/Skylake_X.html

Desktop Quad-core much lower latency than big die Xeons

7-cpu Haswell

Intel i7-4770 (Haswell), 3.4 GHz (Turbo Boost off), 22 nm. RAM: 32 GB (PC3-12800 cl11 cr2).
L1 Data Cache Latency = 4 cycles for simple access via pointer
L1 Data Cache Latency = 5 cycles for access with complex address calculation
L2 Cache Latency = 12 cycles

L3 Cache Latency = 36 cycles (3.4 GHz i7-4770)

L3 Cache Latency = 43 cycles (1.6 GHz E5-2603 v3)

L3 Cache Latency = 58 cycles (core9) - 66 cycles (core5) (3.6 GHz E5-2699 v3 - 18 cores)

RAM Latency = 36 cycles + 57 ns (3.4 GHz i7-4770)
RAM Latency = 62 cycles + 100 ns (3.6 GHz E5-2699 v3 dual)

http://www.7-cpu.com/cpu/Haswell.html

7-cpu Broadwell

Intel i7-6900K (Broadwell), 4.0 GHz, 14 nm. RAM: 32 GB (unknown).
Intel E5-2699 v4 (Broadwell), 3.6 GHz (Turbo Boost), 14 nm. RAM: 256 GB, PC4-2133.
L3 cache = 20 MB, 64 B/line, 20-WAY (i7-6900K)
L3 cache = 55 MB, 64 B/line, 20-WAY (E5-2699 v4)

L1 Data Cache Latency = 4 cycles for simple access via pointer
L1 Data Cache Latency = 5 cycles for access with complex address calculation
L2 Cache Latency = 12 cycles

L3 Cache Latency = 59 cycles (i7-6900K, 4.0 GHz)
L3 Cache Latency = 65 cycles (E5-2699 v4, 3.6 GHz) (60 - 69 cycles on different cores)

RAM Latency = 59 cycles + 46 ns (i7-6900K, 4.0 GHz)
RAM Latency = 65 cycles + 75 ns (E5-2699 v4, 3.6 GHz)

http://7-cpu.com/cpu/Broadwell.html

7-cpu Skylake X

Intel i7-7820X (Skylake X), 8 cores, 4.3 GHz (Turbo Boost), Mesh 2.4 GHz, 14 nm.
RAM: 4x 8 GB DDR4-3400 16-18-18-36.
L2 cache = 1024 KB, 64 B/line, 16-WAY
L3 cache = 11 MB, 64 B/line, 11-WAY

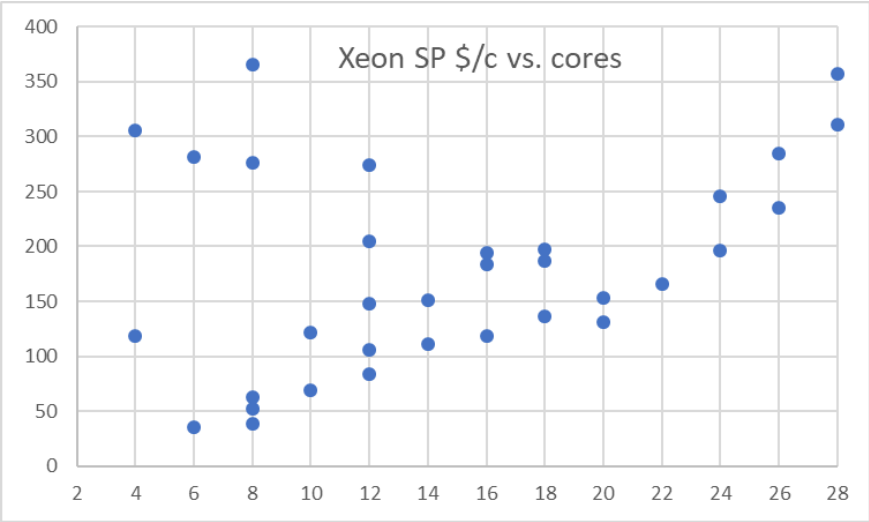
L1 Data Cache Latency = 4 cycles for simple access via pointer
L1 Data Cache Latency = 5 cycles for access with complex address

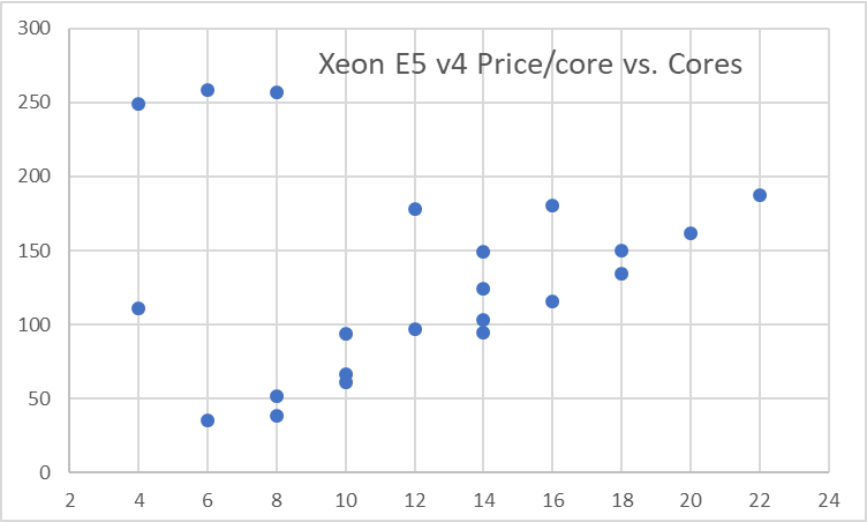
L2 Cache Latency = 14 cycles
L3 Cache Latency = 68 cycles (3.6 GHz)
L3 Cache Latency = 79 cycles (4.3 GHz) (77-81 cycles for different cores)

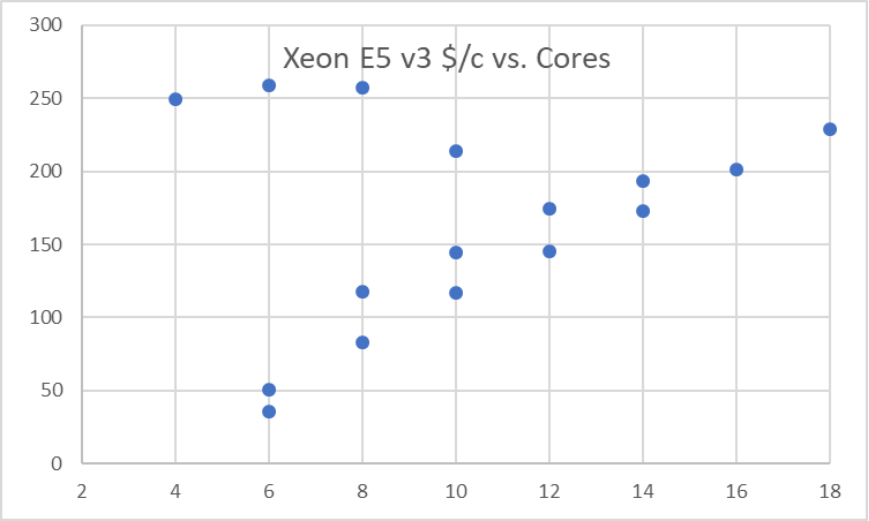
RAM Latency = 79 cycles + 50 ns

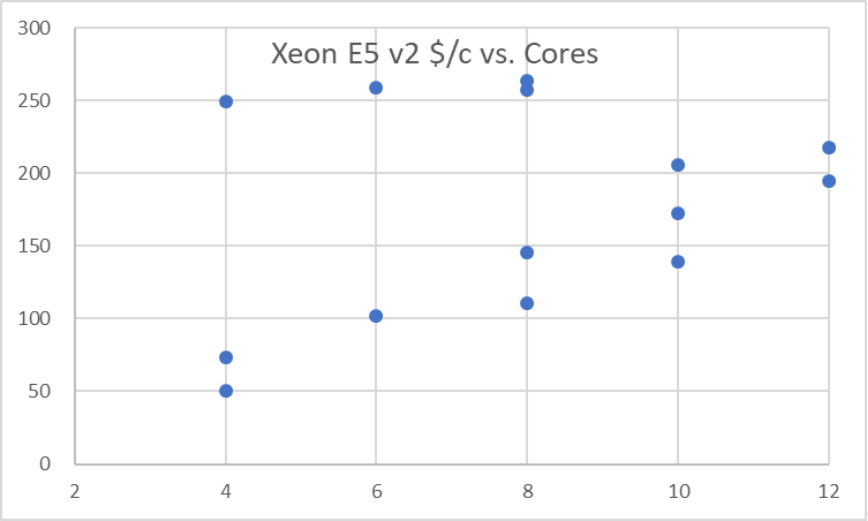
http://7-cpu.com/cpu/Skylake_X.html

Xeon E5 v2, 3, 4 & SP SKUs - cores

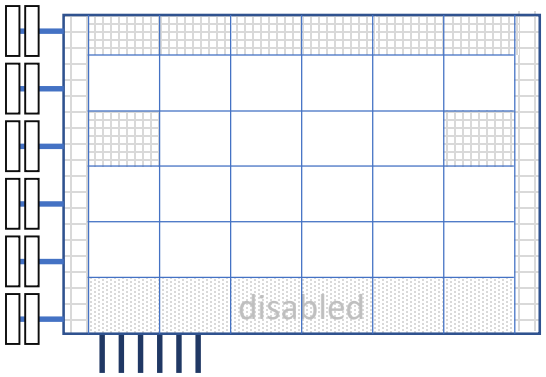






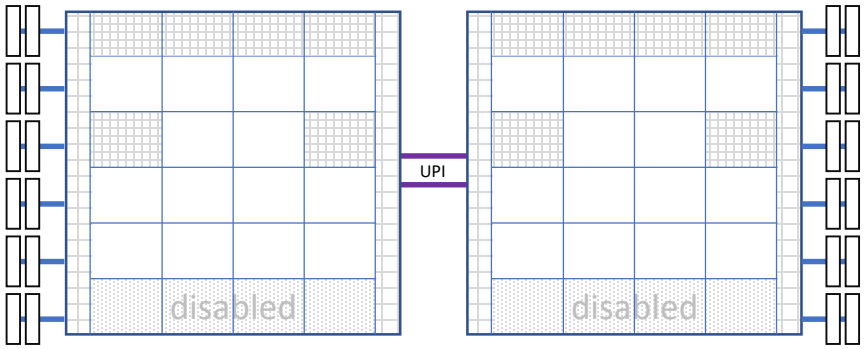


1-2 sockets, by ~equivalent TPS



1 x Xeon Gold 6148
2.4GHz, 20 cores, HT

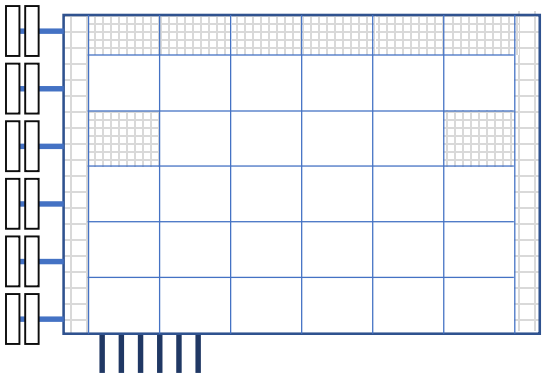
Processor	$3,096
Base + 1 CPU	$4,963
Xeon 6152	$3,655
22c, 2.1GHz	



2 x **Xeon Gold 5120**
2.2GHz, 14 cores per die, HT

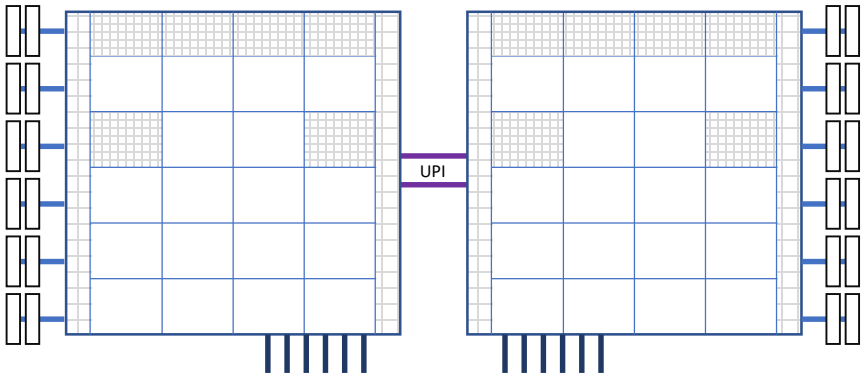
Base system	$1,900
Processor	$1,555 ea.
Base + 2 CPU	$5,010

Equivalent TPS, Alternate



1 x **Xeon Platinum 8180**
2.5GHz, 28 cores, HT

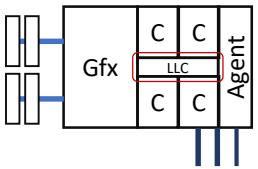
Processor	$10,876 ea.
Base + 1 CPU	$12,756



2 x **Xeon Gold 6140**
2.3GHz, 18 cores per die, HT

Base system	$1,900
Processor	$2,445 ea.
Base + 2 CPU	$6,790

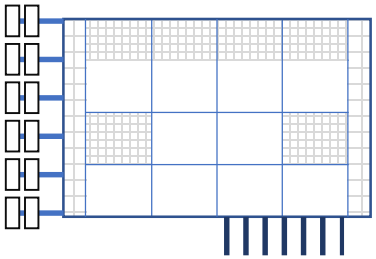
Xeon E3 or Xeon SP?

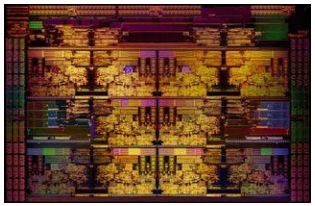


Xeon E3 v6
4.1GHz, 4 cores, HT

2 memory channels
64GB max mem
16 PCI-E + DMI

$284-450





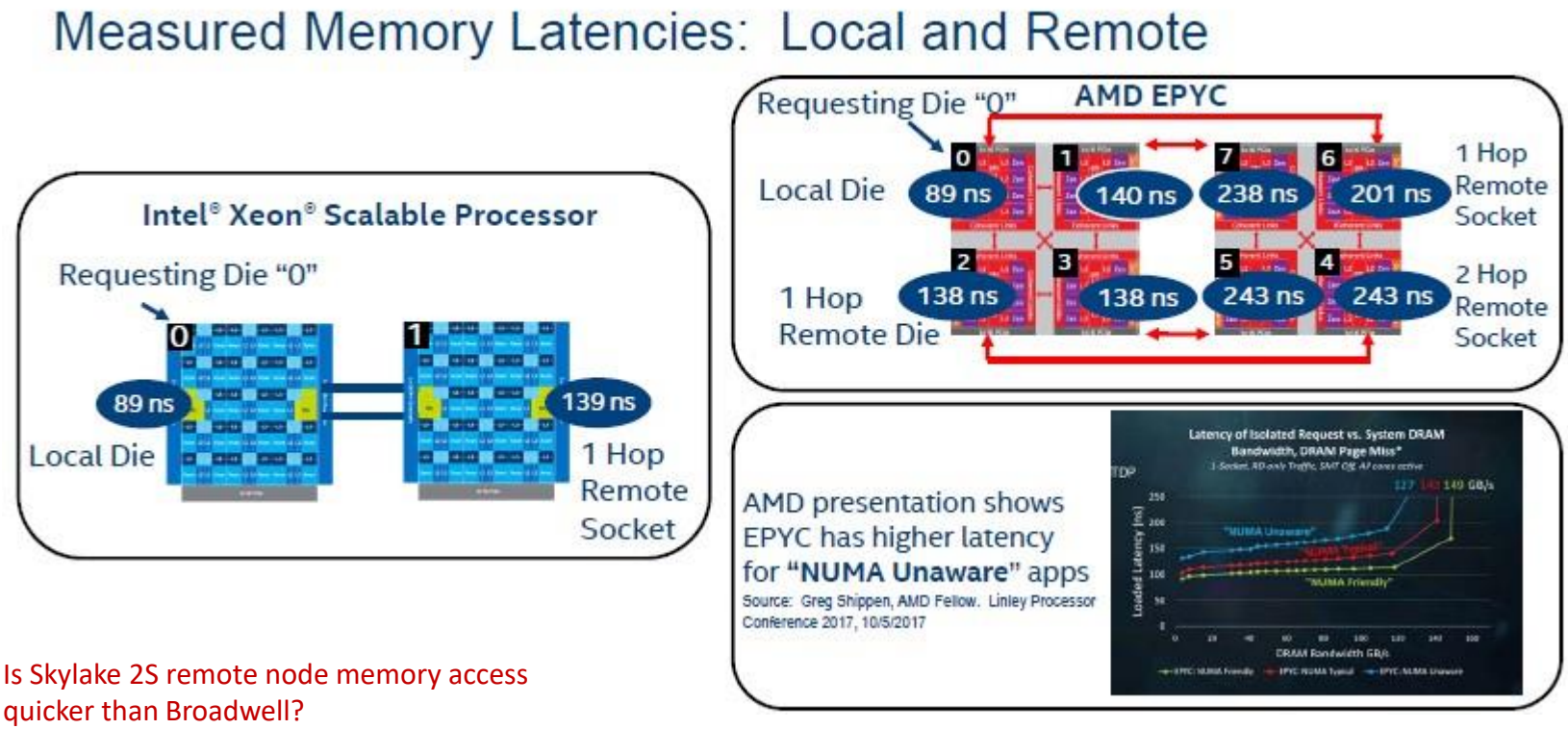
Xeon SP
6-12 cores, HT

6 memory channels
768GB max mem
48 PCI-E + DMI

4116	12c	2.1GHz	$1,002
4114	10c	2.2GHz	$694
3106	8c	1.7GHz	$306
3104	6c	1.7GHz	$213

Low Core Count (LCC)
14 nm process
- ~22.26 mm x ~14.62 mm
- ~325.44 mm² die size

Skylake



Is Skylake 2S remote node memory access quicker than Broadwell?
Or is Broadwell remote node less than 153ns

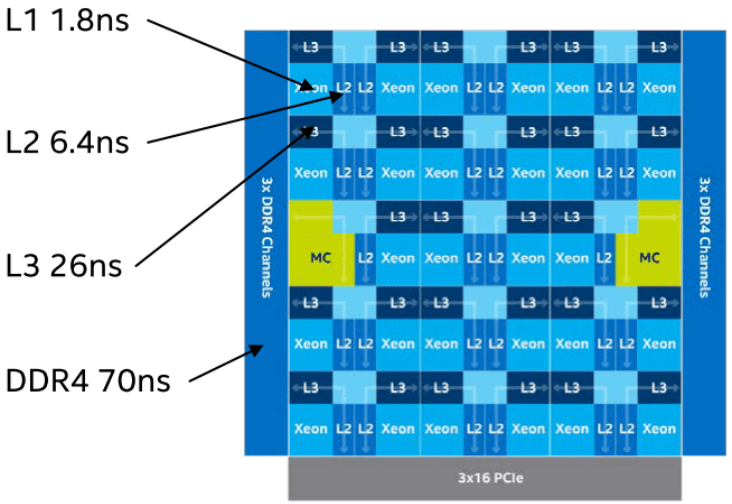
https://www.nextplatform.com/2017/11/27/intel-stacks-xeons-amd-epyc-systems/

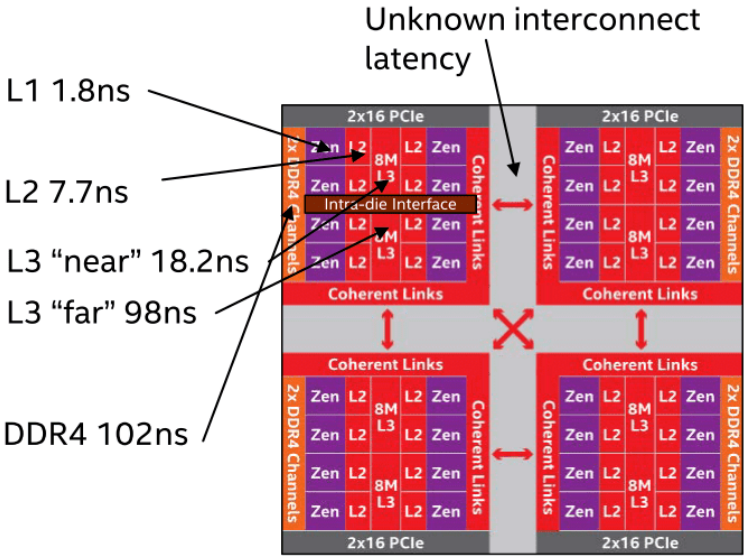
Memory and Cache

Content Under Embargo Until 9:15 AM PST July 11, 2017

Memory and Cache Latency Comparisons

Implementation Matters

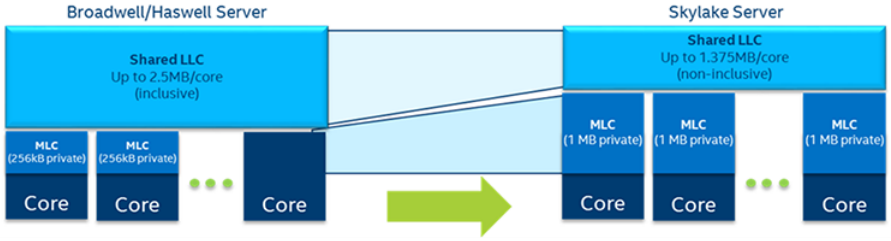


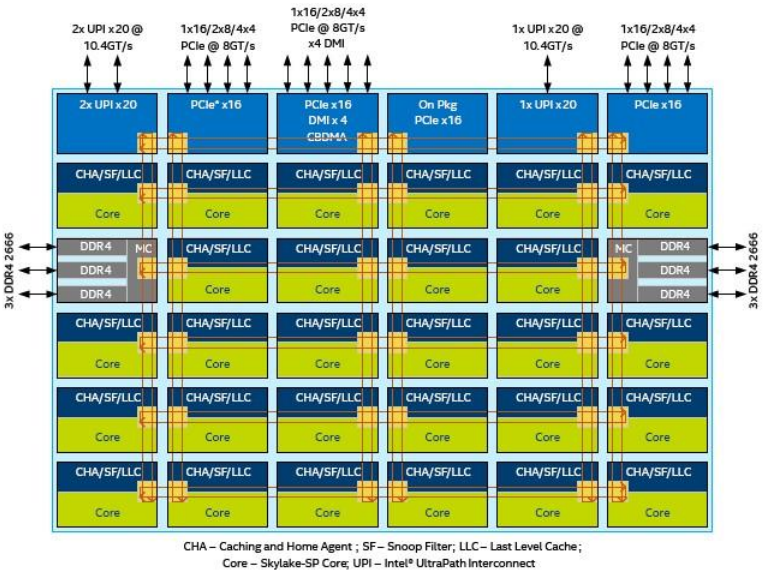


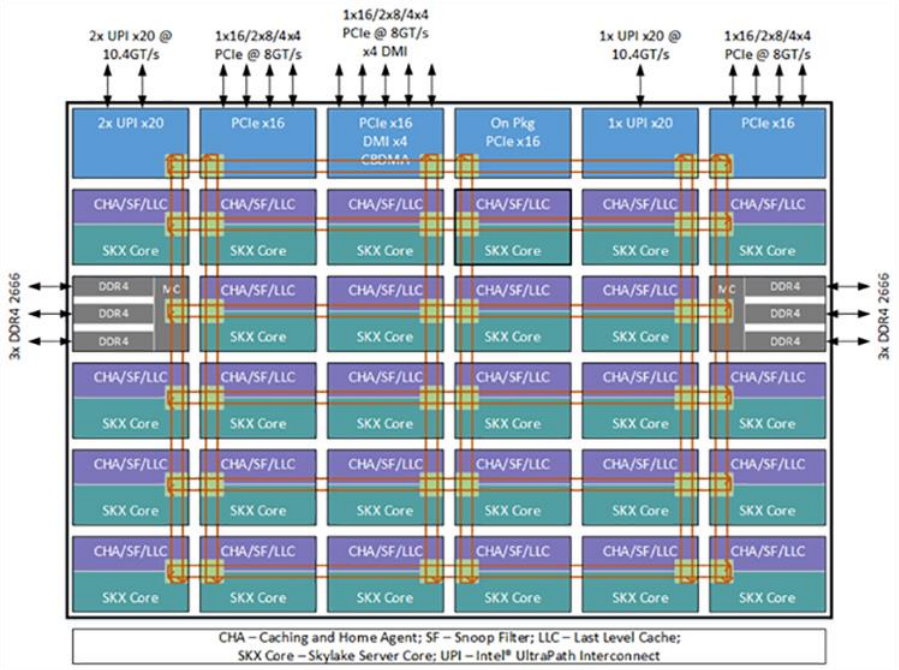
Lower is better

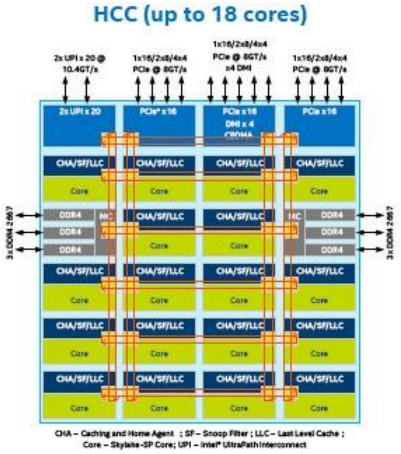
L3: Intel up 73% lower latency than AMD
DDR: Intel up to 31% lower latency

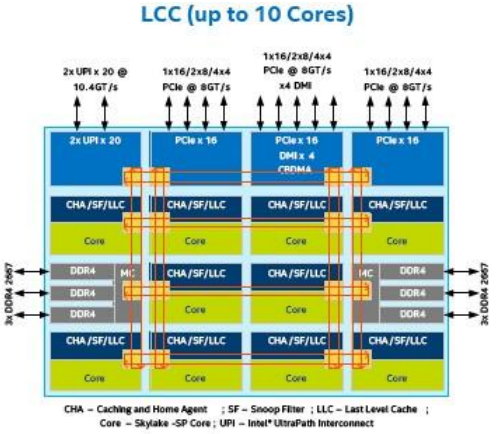
Skylake SP diagrams



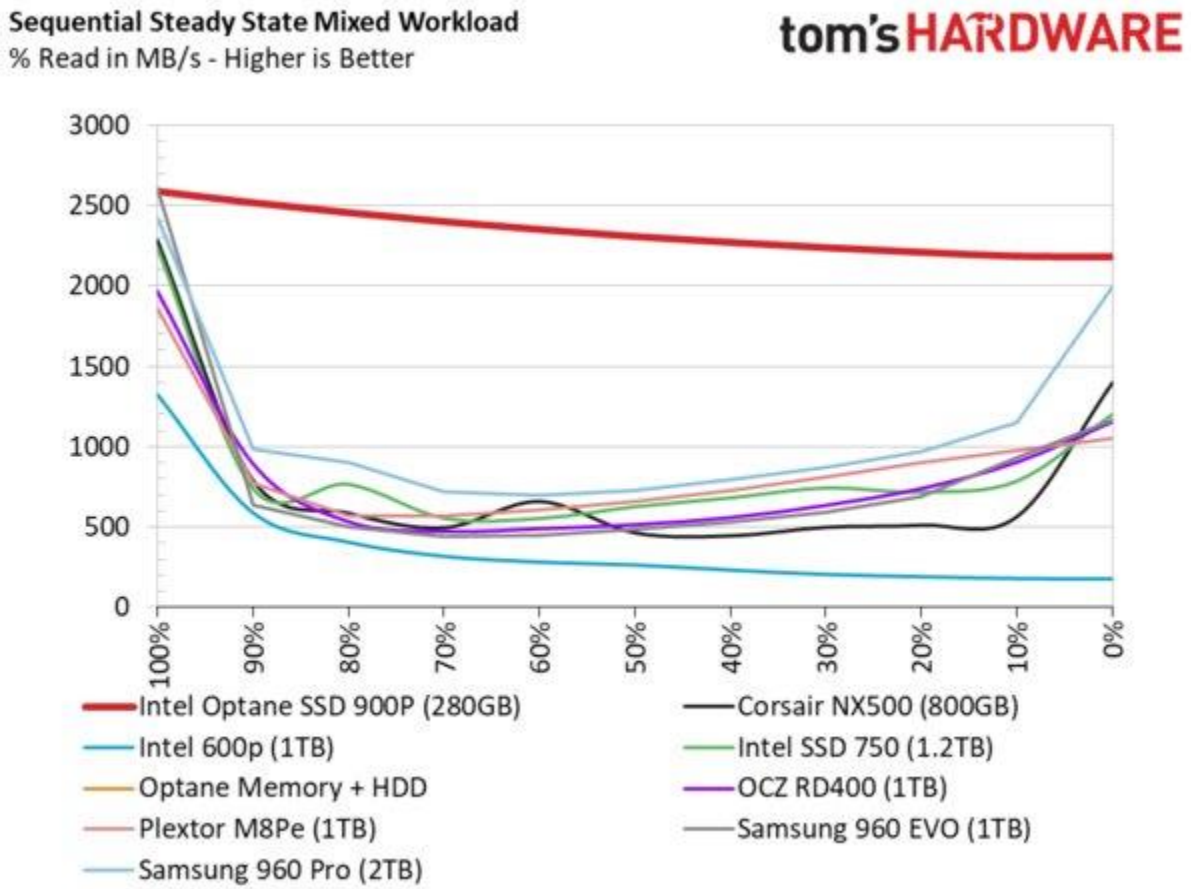








Intel 3D XPoint/Optane vs NAND

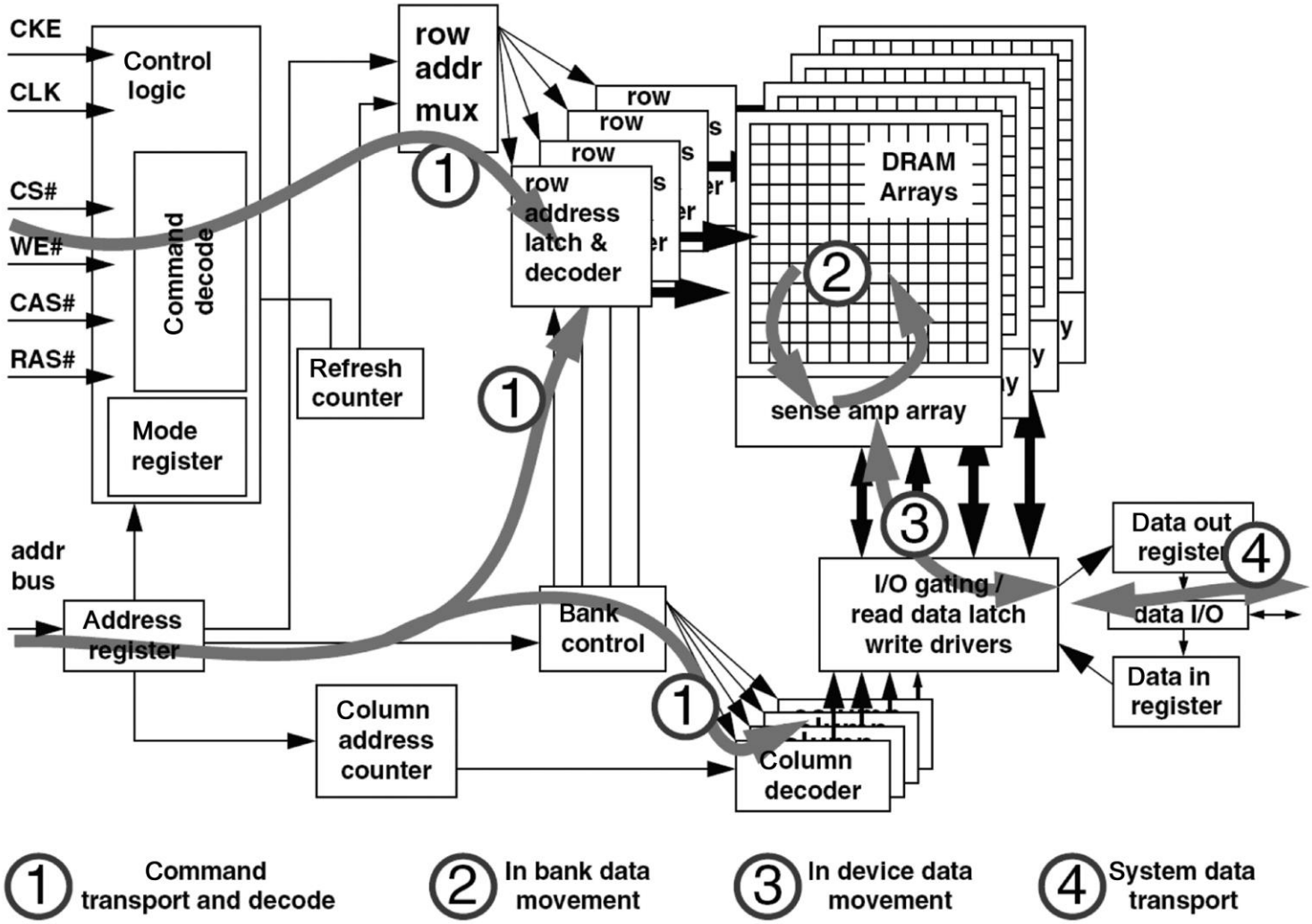


http://www.tomshardware.com/reviews/intel-optane-ssd-900p-3d-xpoint,5292-2.html

Memory Bandwidth

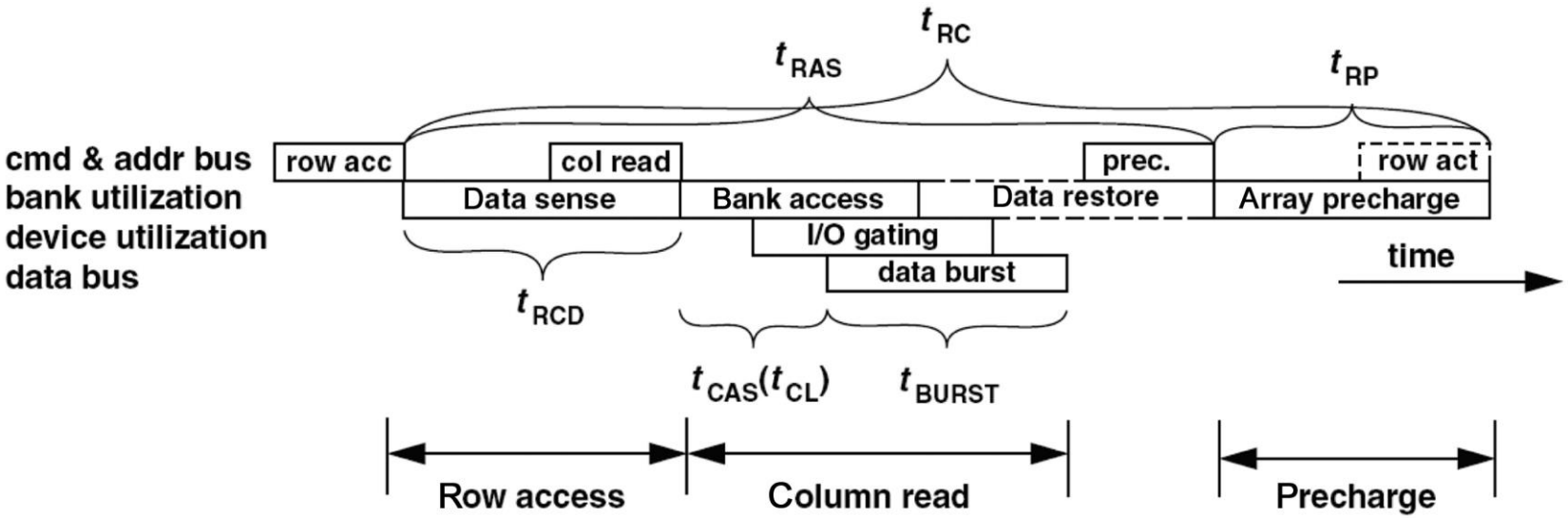
DDR4

from “Bounding Worst-Case DRAM Performance on Multicore Processors”

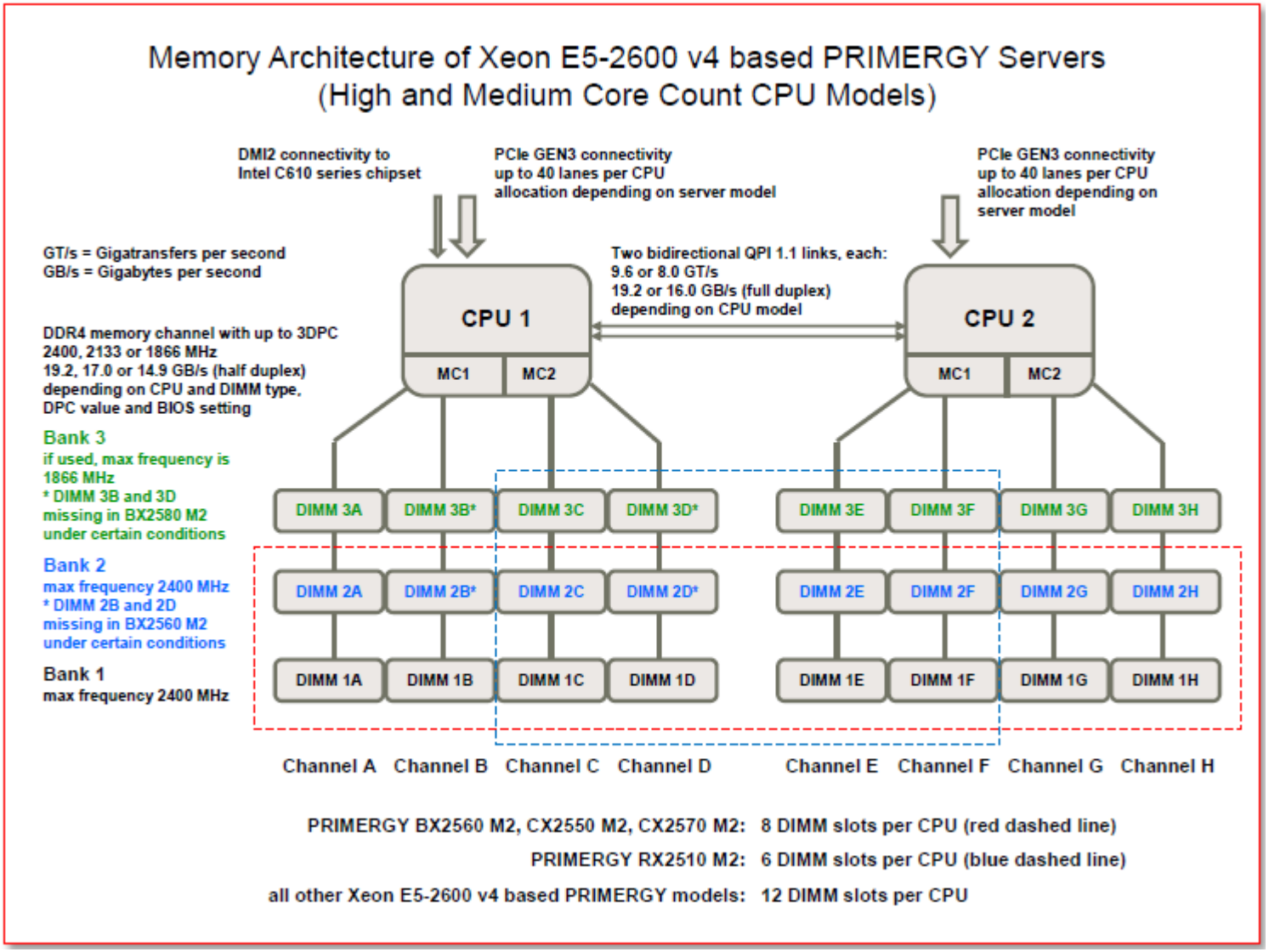


http://central.oak.go.kr/journallist/journaldetail.do?article_seq=11983&tabname=abst Ding Yiqiang, Wu Lan, Zhang Wei

ditto



http://central.oak.go.kr/journallist/journaldetail.do?article_seq=11983&tabname=abst Ding Yiqiang, Wu Lan, Zhang Wei



https://sp.ts.fujitsu.com/dmsp/Publications/public/wp-broadwell-ep-memory-performance-ww-en.pdf

Micron 8Gb DDR4

Table 1: Key Timing Parameters

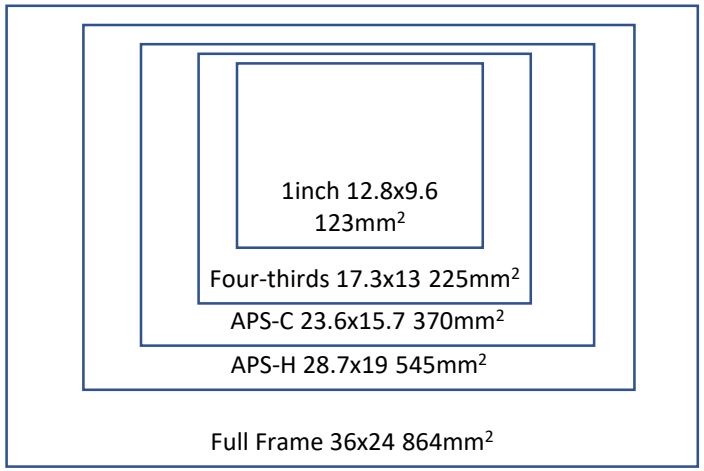
Speed Grade	Data Rate (MT/s)	Target CL-^{t}RCD-^{t}RP	CL (ns)	^{t}RCD (ns)	^{t}RP (ns)
-062E[6]	3200	22-22-22	13.75	13.75	13.75
-068[5]	2933	21-21-21	14.32	14.32	14.32
-075[4]	2666	19-19-19	14.25	14.25	14.25
-075E[4]	2666	18-18-18	13.50	13.50	13.50

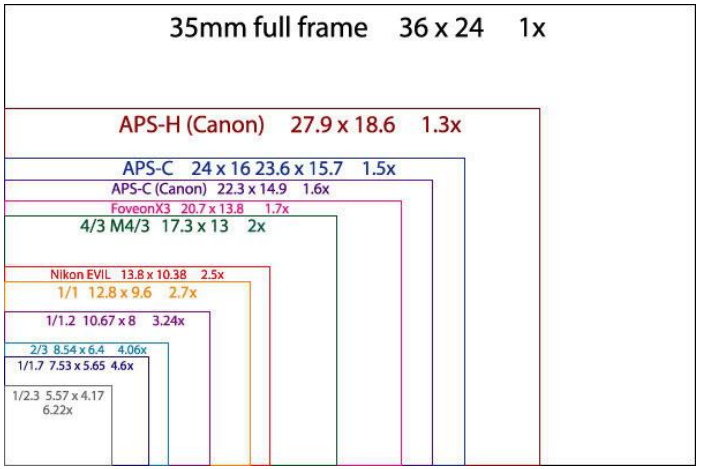
Table 1: Key Timing Parameters (Continued)

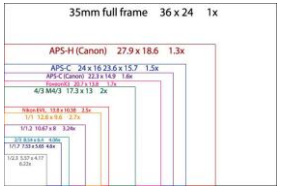
Speed Grade	Data Rate (MT/s)	Target CL-^{t}RCD-^{t}RP	CL (ns)	^{t}RCD (ns)	^{t}RP (ns)
-083[3]	2400	17-17-17	14.16	14.16	14.16
-083E[3]	2400	16-16-16	13.32	13.32	13.32
-093E[2]	2133	15-15-15	14.06	14.06	14.06
-107E[1]	1866	13-13-13	13.92	13.92	13.92

Notes:
1. Backward compatible to 1600, CL = 11.
2. Backward compatible to 1600, CL = 11 and 1866, CL = 13.
3. Backward compatible to 1600, CL = 11; 1866, CL = 13; and 2133, CL = 15.
4. Backward compatible to 1600, CL = 11; 1866, CL = 13; 2133, CL = 15; and 2400, CL = 17.
5. Backward compatible to 1600, CL = 11; 1866, CL = 13; 2133, CL = 15; 2400, CL = 17; and 2666, CL = 19. Speed offering may have restricted availability.
6. Backward compatible to 1600, CL = 11; 1866, CL = 13; 2133, CL = 15; 2400, CL = 17; 2666, CL = 19. Speed offering may have restricted availability.

Camera Sensor Sizes







Scale 10mm to 1in

Additional work

- Matching 1S & 2S systems,
  - Either Xeon E5 v4 (or 3) or Xeon SP
  - Same cores/socket or 1S 14c vs 2S 10c, (1.4:2x1)
- Benchmark artificial & real transaction workload
  - verify actual versus predicted 1S - 2S scaling
- Demonstrate NUMA effect
  - Concentrated inserts to table with identity key
  - versus distributed key
- Deploy Strategy, evaluate efficiency savings
- Whitepaper – w/good quantitative analysis