Scaling through more cores

From single to multi core

by Thomas Walther
Seminar on 30.11.2015

Index

Index

Scaling with single core until 2005

Moore's law	–	Transistors are doubled every 12 to 24 month
Smaller structures	–	New manufacturing technologies
Higher frequencies	–	Smaller structures need less power

Index

Problems and barriers: Problems

Higher frequencies	–	Needed for more computing power
Higher voltage	–	Needed for higher frequencies
Heat production	–	High frequencies and high voltage result in higher power consumption and dissipation (waste heat)

Problems and barriers: Barriers

Power consumption	–	Frequency and voltage directly influence the consumption: $P = a * C * V^2 * f$
Hot spots	–	Smaller structures result in smaller hot spots, that are more difficult to cool
Critical Point 2004	–	Pentium 4 with ~4 Ghz (air cooled) marked the line for the next years

Note: P = power(Watts), V = voltage(Volts), f = frequency(cycles/sec), C = capacitance(Farads), a = Coulombs/Volt

Index

Solution through more cores

Since 2006	–	Intel accepted the end of single cores and introduced the first Dual Cores with 1.5 to 2.33 Ghz
More Cores and HT	–	Intel combined more cores on one DIE with smaller structures and a customized architecture and later additionally the cheaper Hyper Threading
Heat production	–	Lower work load per core reduces the power consumption and heat production in total

Index

1. Introduction

2. Scaling with single core until 2005
   Problems and barriers

3. **Solution through more cores**
   Current standard
   Multi core usage
   Amdahl's law

4. Problems and barriers
   Hot spots
   DIE size

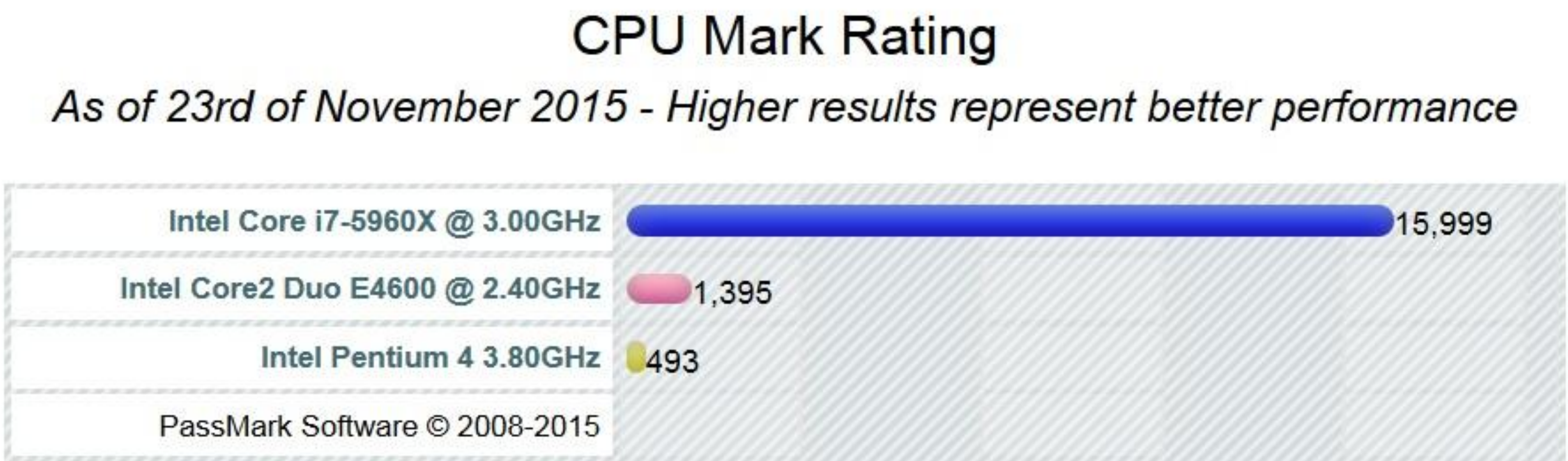
5. Fabrication and approaches for new technologies
   New materials for higher frequencies

6. Conclusion

Current standard

Consumer PC	–	2 to 8 cores with a total of 16 threads
Handheld	–	2 to 8 cores
HPC in the Top 500	–	From thousands up to 3.12 million cores

Current standard



Graphic 1, s. Attachment

First multi cores	–	In comparison to Pentium 4 the Core2 Duo processor is ~2.8 times faster
Multi cores today	–	In comparison to Pentium 4 one of the newest processors with 8 cores and 16 threads is ~32.5 times faster

Current standard

	Intel Pentium 4 3.80GHz	Intel Core2 Duo E4600 @ 2.40GHz	Intel Core i7-5960X @ 3.00GHz
Socket Type	NA[2]	LGA775	LGA2011-v3
CPU Class	Desktop	Desktop	Desktop
Clockspeed	3.8 GHz	2.4 GHz	3.0 GHz
Turbo Speed	Not Supported	Not Supported	Up to 3.5 GHz
# of Physical Cores	1 (2 logical cores per physical)	2	8 (2 logical cores per physical)
Max TDP	65W	65W	140W
First Seen on Chart	Q4 2008	Q4 2008	Q2 2014
# of Samples	52	406	371
Single Thread Rating	824[3]	888	1993
CPU Mark	**493**	**1395**	**15999**

[1] - *Last seen price from our affiliates* ***NewEgg.com*** *&* ***Amazon.com****.*

[2] - *Information not available. Do you know?* ***Notify Us****.*

[3] - *Single thread rating may be higher than the overall rating, thread performance is just one* ***component*** *of the CPU Mark.*

Graphic 2, s. Attachment

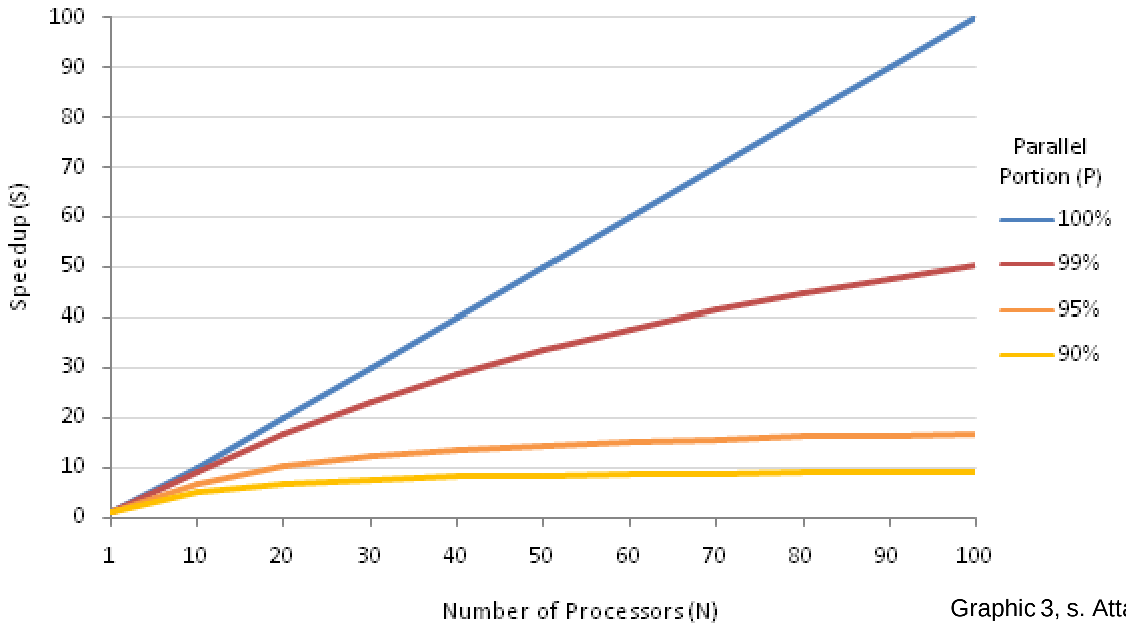
Index

Multi core usage

Consumer/Handheld	–	Only a few programs that use more than 2 cores, especially games or graphic programs
HPC usage	–	Programs are limited by hardware – open potential!
Parallelization	–	The hardware power needs a optimized parallelization to achieve its full potential
Potential limited	–	Amdahl’s law shows the limits

Index

Amdahl's law

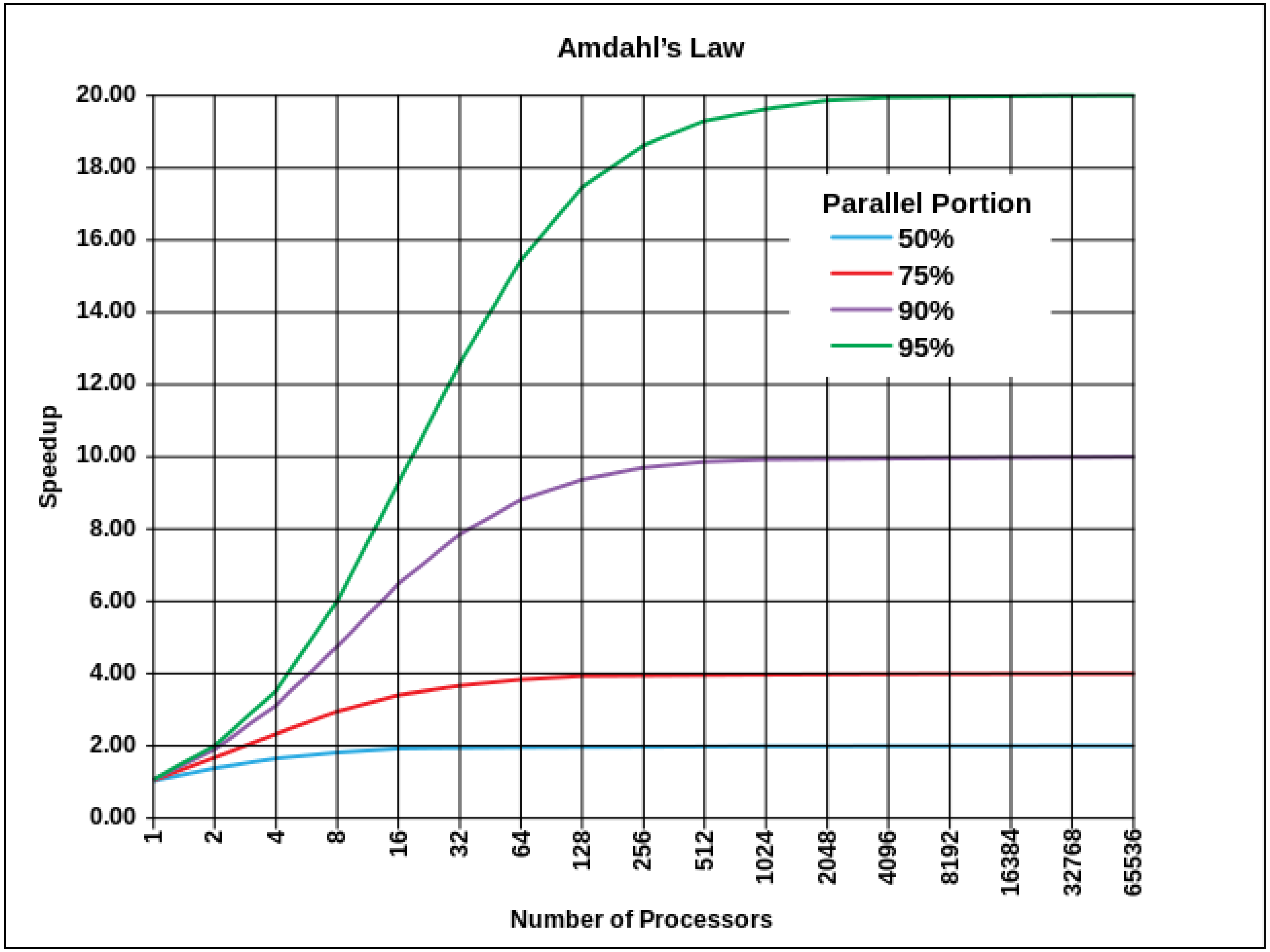


Graphic 3, s. Attachment

Amdahl's law $$\eta_S = \frac{T}{t_S + t_{O(n_P)} + \frac{t_P}{n_P}} \leq \frac{T}{T - t_P}$$

Note: *ns* = speedup, T = running time(RT), ts = serial RT, tp = parallel RT, np = cores, to(np) = synchron time

Amdahl's law



Graphic 4, s. Attachment

Amdahl’s law

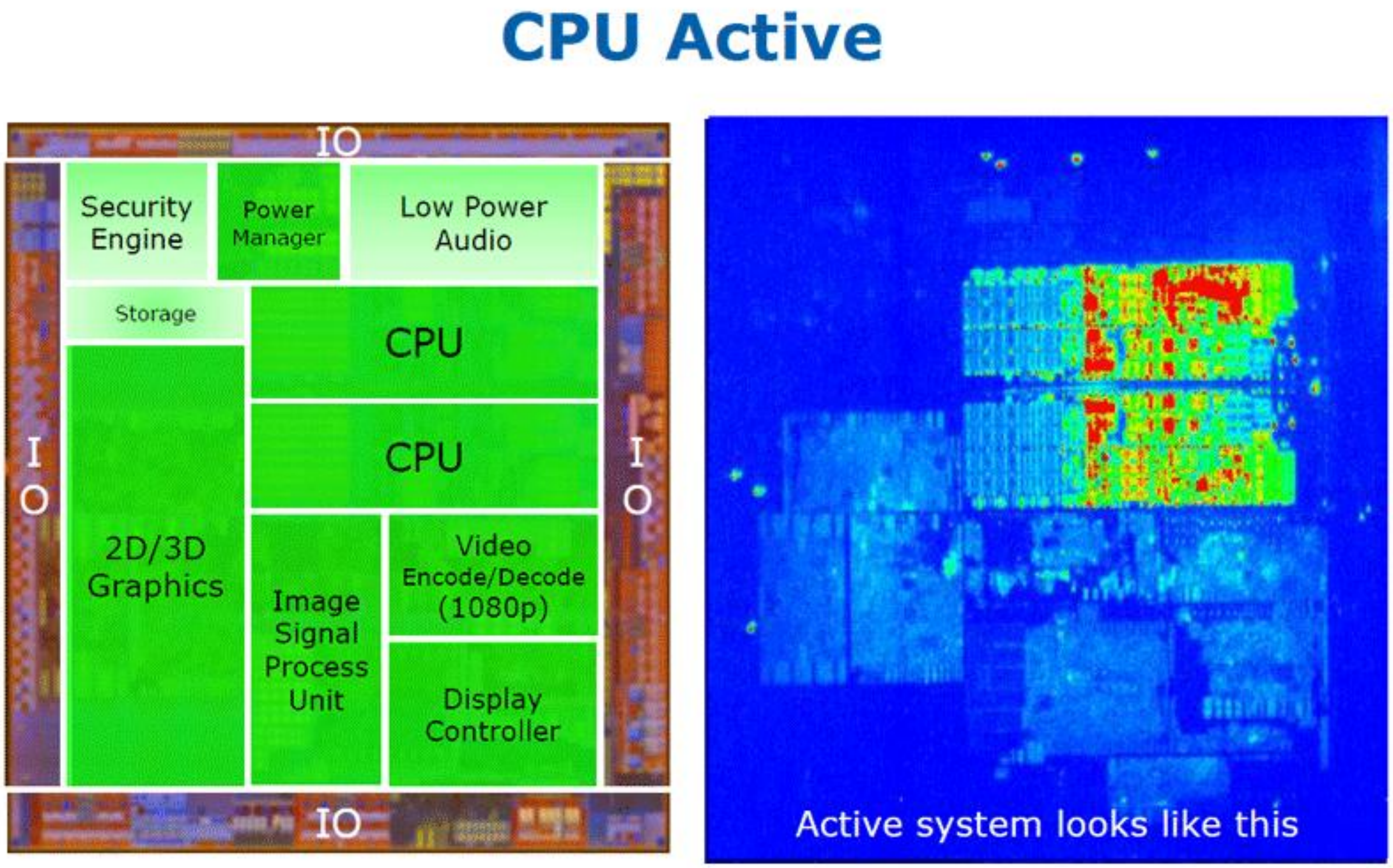
Parallelization	–	Not every problem can be solved parallel if it is not possible to split it up and compute the problems separately
Time problem	–	A program running time can never be reduced below its serial components runtime
Tianhe-2	–	With 3.12 Million cores the computing power with >33 PFLOPS has a huge potential, but nearly no program can use it at once

Index

Hot spots

Small structures	–	Hot spots get more intense the smaller the structures get
High frequencies	–	Also multi cores have the weakness when one core clocks higher, that one small point gets hot
Hot Spot cooling	–	High temperatures in small areas are still hard to cool and the different cooling methods limit the possible waste heat

Hot spots



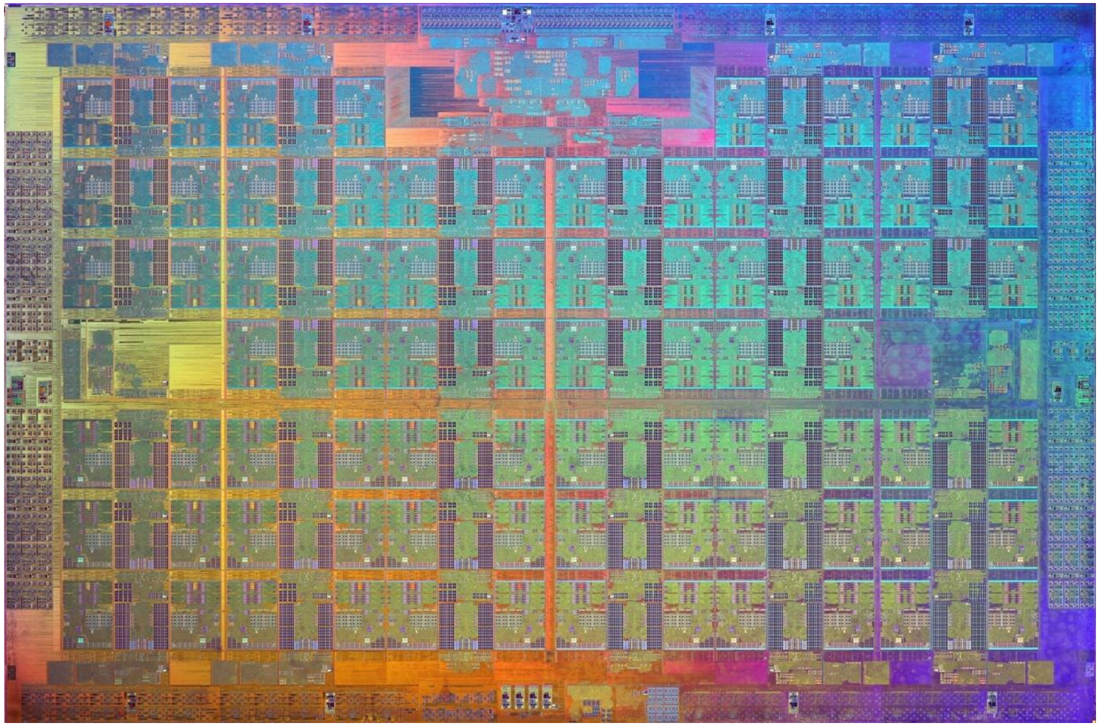
Graphic 5, s. Attachment

Index

DIE size

Limited space on DIE	–	Space needed for the CPU, cache, register and today often the GPU
Smaller structures	–	More space for more technology but smaller hot spots
Knights Landing	–	Even with very low frequencies the heat generation will explode with too many cores (Up to 72(76) cores with a TDP of 200W)

DIE size: Intel Xeon PHI – Knights Landing



Graphic 6, s. Attachment

DIE size – 700 mm² produced with 14 nm Lithography

Index

Fabrication and approaches for new technologies

New architectures	–	Reprogrammable simple CPU parts for daily tasks to use the main CPU part on DIE for one special task
New cooling materials	–	Directly distributed between the components like a viscous mass to spread the heat

Index

New materials for higher frequencies

Semiconductor materials – Possible successor of silicon could be indium, germanium and gallium arsenide

Reduced voltage – This alternative materials can run at 0.5V while silicon needs a voltage around 1.1V

Higher frequencies – Reduced voltage requirement leads to higher frequencies to use the full potential

Index

Conclusion

Problems and barriers	–	Power consumption always exists and needs permanent optimization through new technologies
Parallelization	–	Developer have to learn to program and optimize their software for parallelization
New pathes	–	Specialized Hardware with its own software and code is a possible way for more speed

Sources

http://www.extremetech.com/extreme/166413-post-post-pc-the-new-materials-tech-and-cpu-designs-that-will-revive-overclocking-and-enthusiast-computing

https://de.wikipedia.org/wiki/Amdahlsches_Gesetz

https://de.wikipedia.org/wiki/Mooresches_Gesetz

https://software.intel.com/en-us/blogs/2009/06/29/why-p-scales-as-cv2f-is-so-obvious

https://en.wikipedia.org/wiki/Tianhe-2

https://en.wikipedia.org/wiki/List_of_CPU_power_dissipation_figures#Desktop_processors

http://lwn.net/Articles/483889/

http://physics.stackexchange.com/questions/34766/how-does-power-consumption-vary-with-the-processor-frequency-in-a-typical-comput

http://superuser.com/questions/988453/why-multi-core-processors-producing-less-heat

http://superuser.com/questions/543702/why-are-newer-generations-of-processors-faster-at-the-same-clock-speed/906227#906227

Graphics:
[1] , [2]: http://www.cpubenchmark.net/compare.php?cmp[]=1081&cmp[]=937&cmp[]=2332

[3] : http://blogs.dnvgl.com/utilityofthefuture/wp-content/uploads/2011/12/Amdahls-law1.gif

[4] : https://upload.wikimedia.org/wikipedia/commons/thumb/e/ea/AmdahlsLaw.svg/800px-AmdahlsLaw.svg.png

[5]: http://www.extremetech.com/wp-content/uploads/2013/09/big_soc-cpu2-e1379614854509-640x403.png

[6]: http://pics.computerbase.de/6/9/0/1/0/1-1260.jpg