Norwegian University of Science and Technology

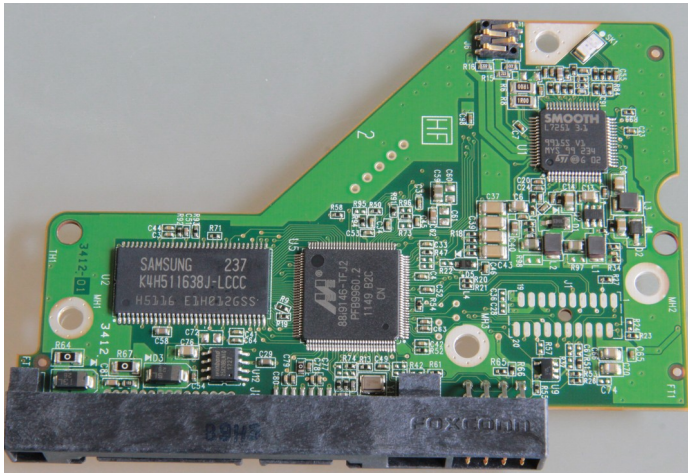
Why parallel computing?

Bart Iver van Blokland

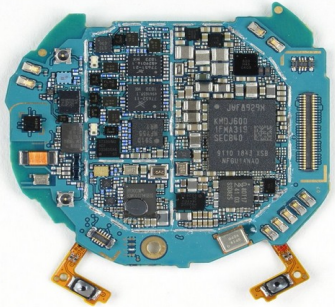
Parallel processors are everywhere..

Product Name	Launch Date	Total Cores	Processor Base Frequency	Cache	TDP
Intel® Quark™ Microcontroller D2000	Q3'15	1	32 MHz	0 KB	
Intel® Quark™ SE C1000 Microcontroller	Q4'15	1	32 MHz	8 KB	
Intel® Quark™ Microcontroller D1000	Q3'15	1	33 MHz	0 KB	0.025 W
Intel® Quark™ SoC X1001	Q2'14	1	400 MHz	16 KB	2.2 W
Intel® Quark™ SoC X1011	Q2'14	1	400 MHz	16 KB	2.2 W
Intel® Quark™ SoC X1020	Q2'14	1	400 MHz	16 KB	2.2 W
Intel® Quark™ SoC X1021	Q2'14	1	400 MHz	16 KB	2.2 W
Intel® Quark™ SoC X1021D	Q2'14	1	400 MHz	16 KB	2.2 W
Intel® Quark™ SoC X1010	Q1'14	1	400 MHz	16 KB	2.2 W
Intel® Quark™ SoC X1020D	Q1'14	1	400 MHz	16 KB	2.2 W
Intel Atom® Processor E3815	Q4'13	1	1.46 GHz	512 KB L2 Cache	5 W
Intel® Quark™ SoC X1000	Q4'13	1	400 MHz	16 KB	2.2 W
Intel® Celeron® Processor G470	Q2'13	1	2.00 GHz	1.5 MB Intel® Smart Cache	35 W
Intel® Celeron® Processor 927UE	Q1'13	1	1.50 GHz	1 MB Intel® Smart Cache	17 W

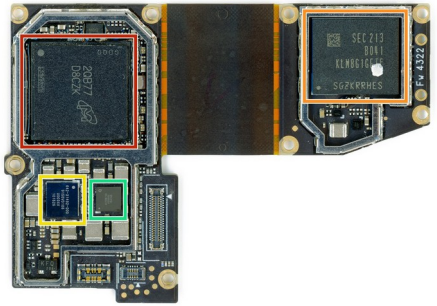
A



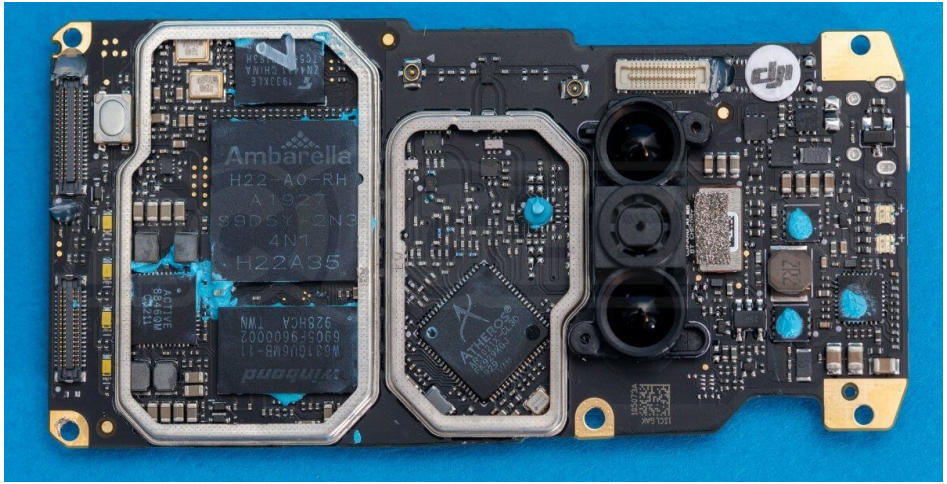
B



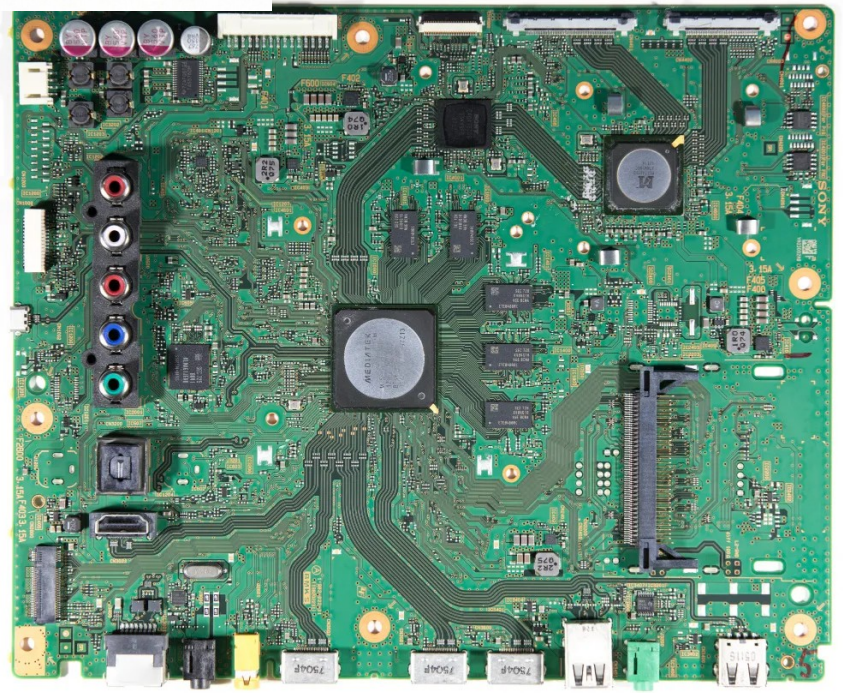
C



D

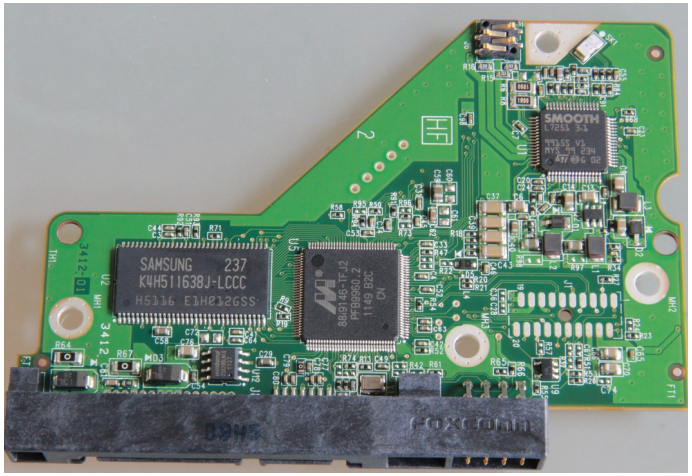


E

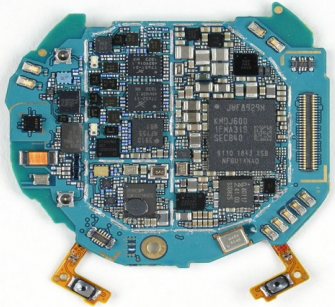


NTNU | Norwegian University of Science and Technology

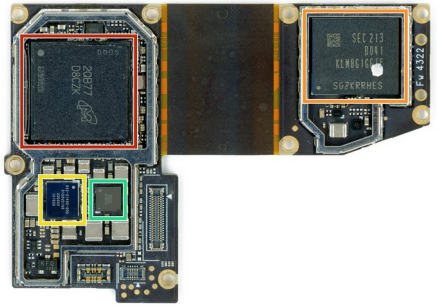
A



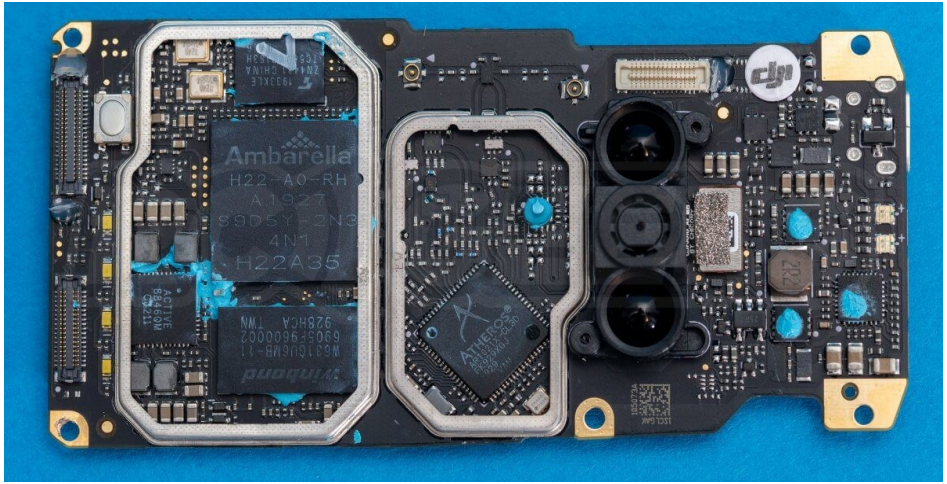
B



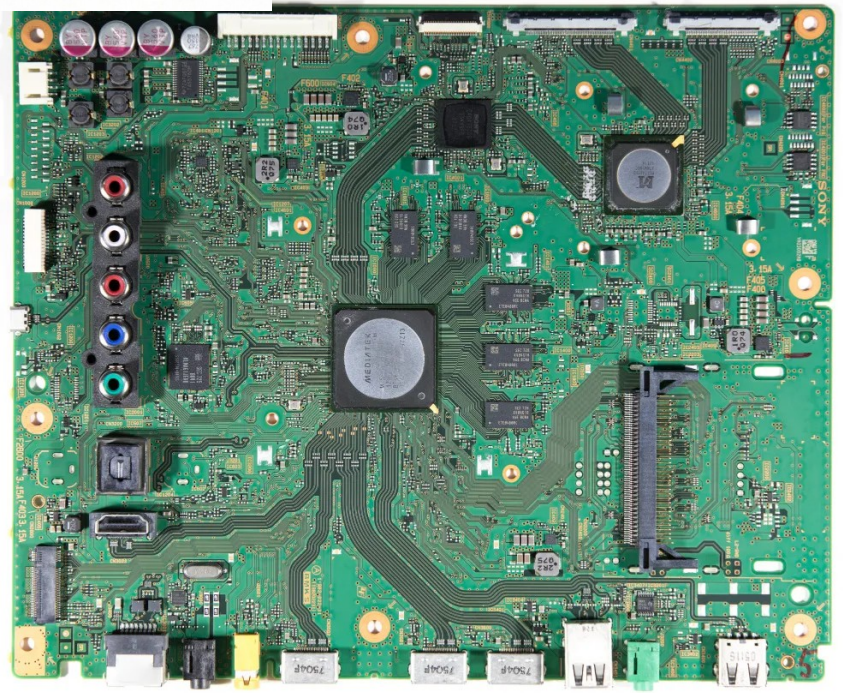
C



D



E

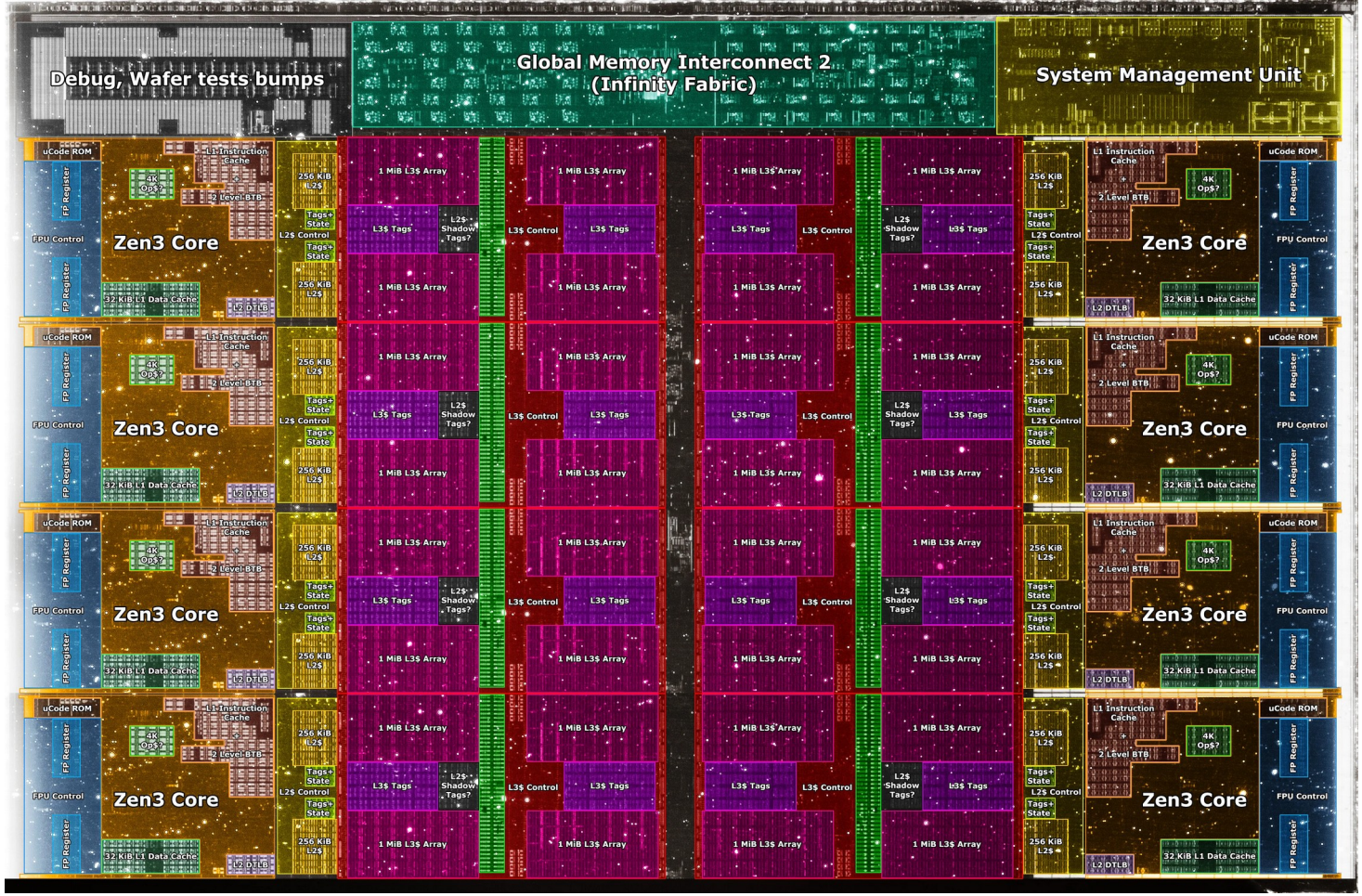


NTNU | Norwegian University of Science and Technology

Microchip

Why parallel computing?

- **Multicore processors are ubiquitous**
- Needed to fully utilise all cores of a processor
- Only way for chip manufacturers to improve performance
- Some processors are practically useless without it
- Problem is too large to fit one machine
- Can even be useful on single core machines

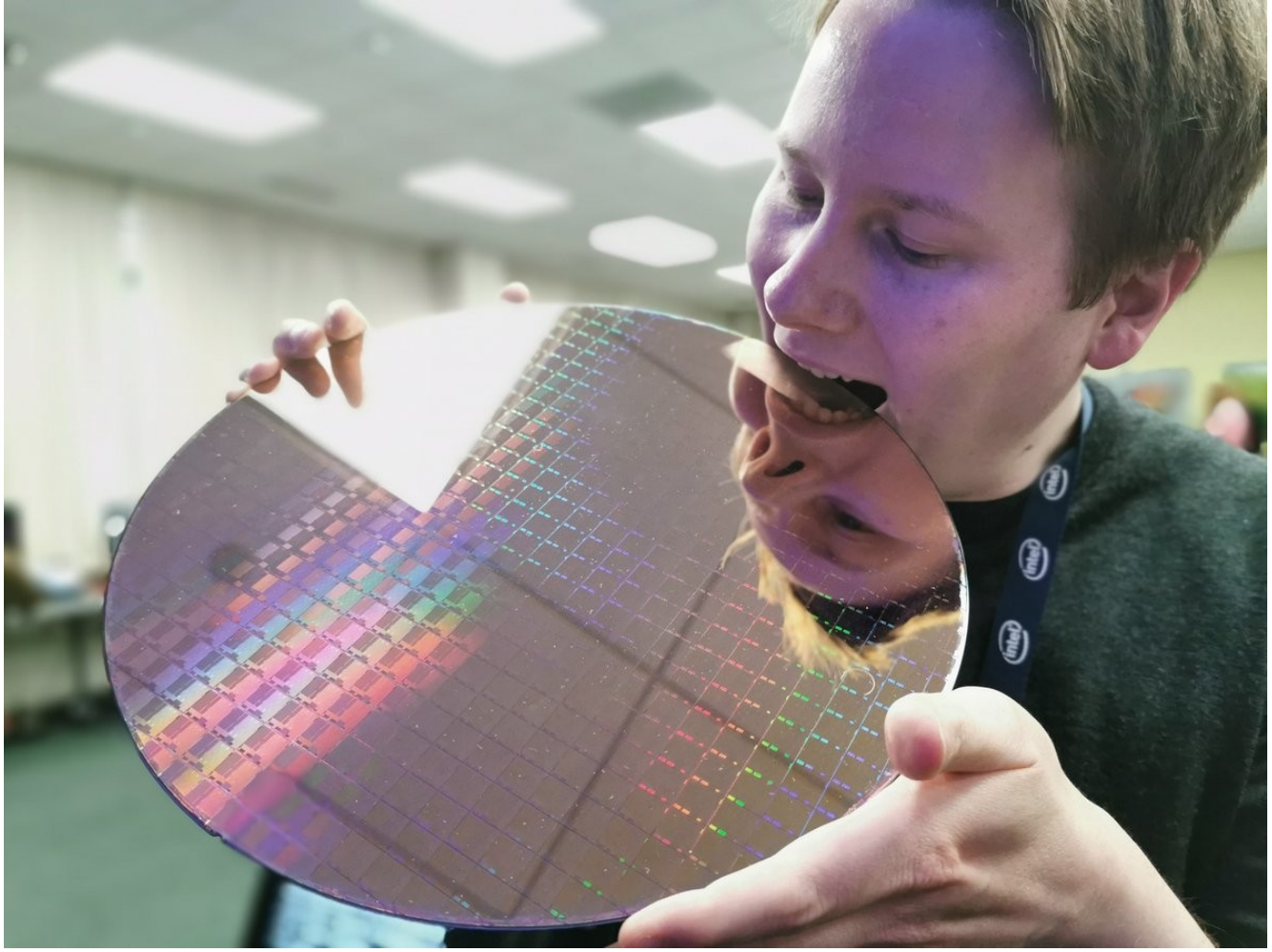

Debug, Wafer tests bumps
Global Memory Interconnect 2
(Infinity Fabric)
System Management Unit
Zen3 Core
uCode ROM
FP Register
FPU Control
L1 Instruction Cache
4K Ops?
2 Level BTB
32 KiB L1 Data Cache
L2 DTLB
256 KiB L2$
Tags+ State
L2$ Control
1 MiB L3$ Array
L3$ Tags
L2$ Shadow Tags?
L3$ Control

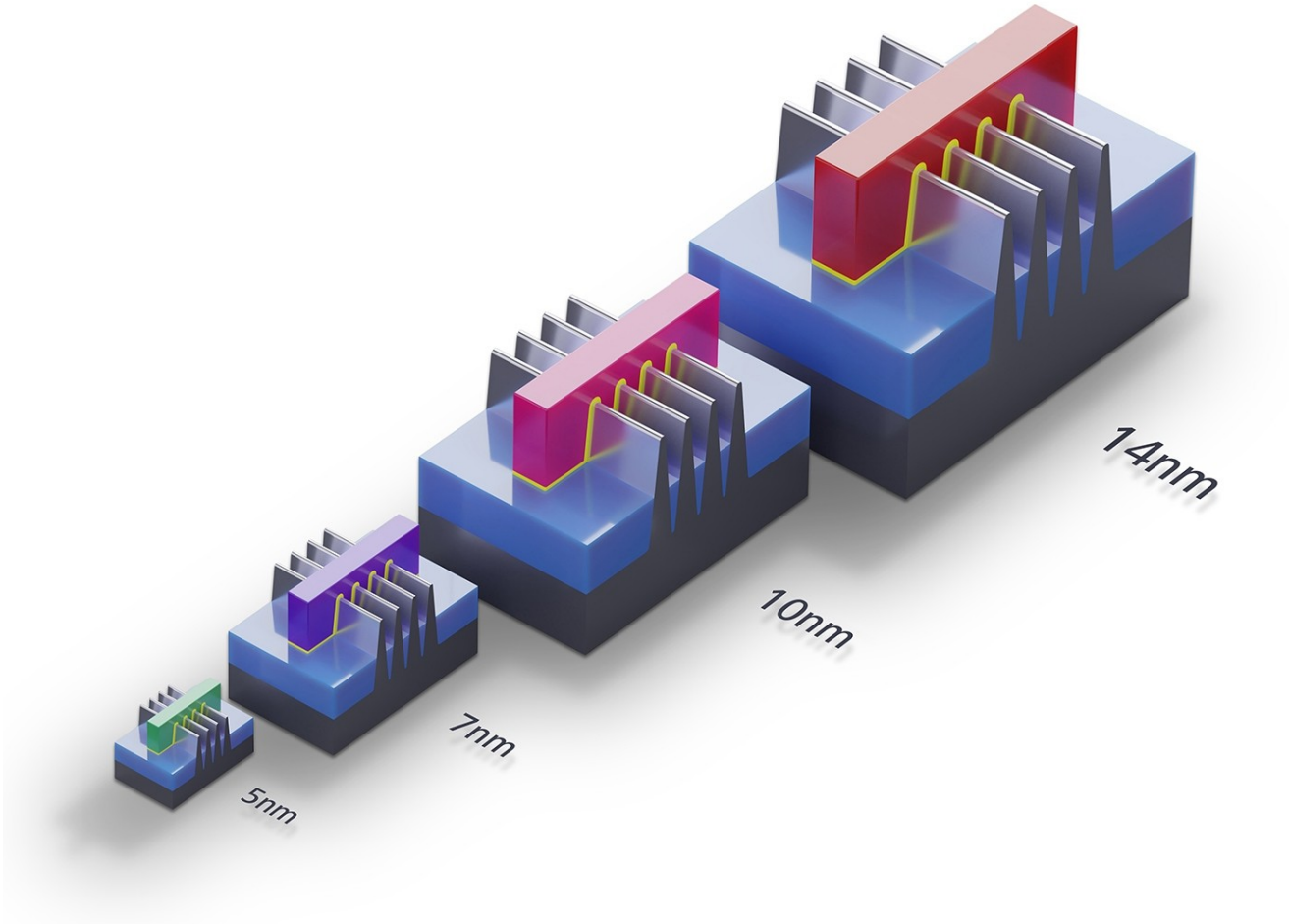
Why parallel computing?

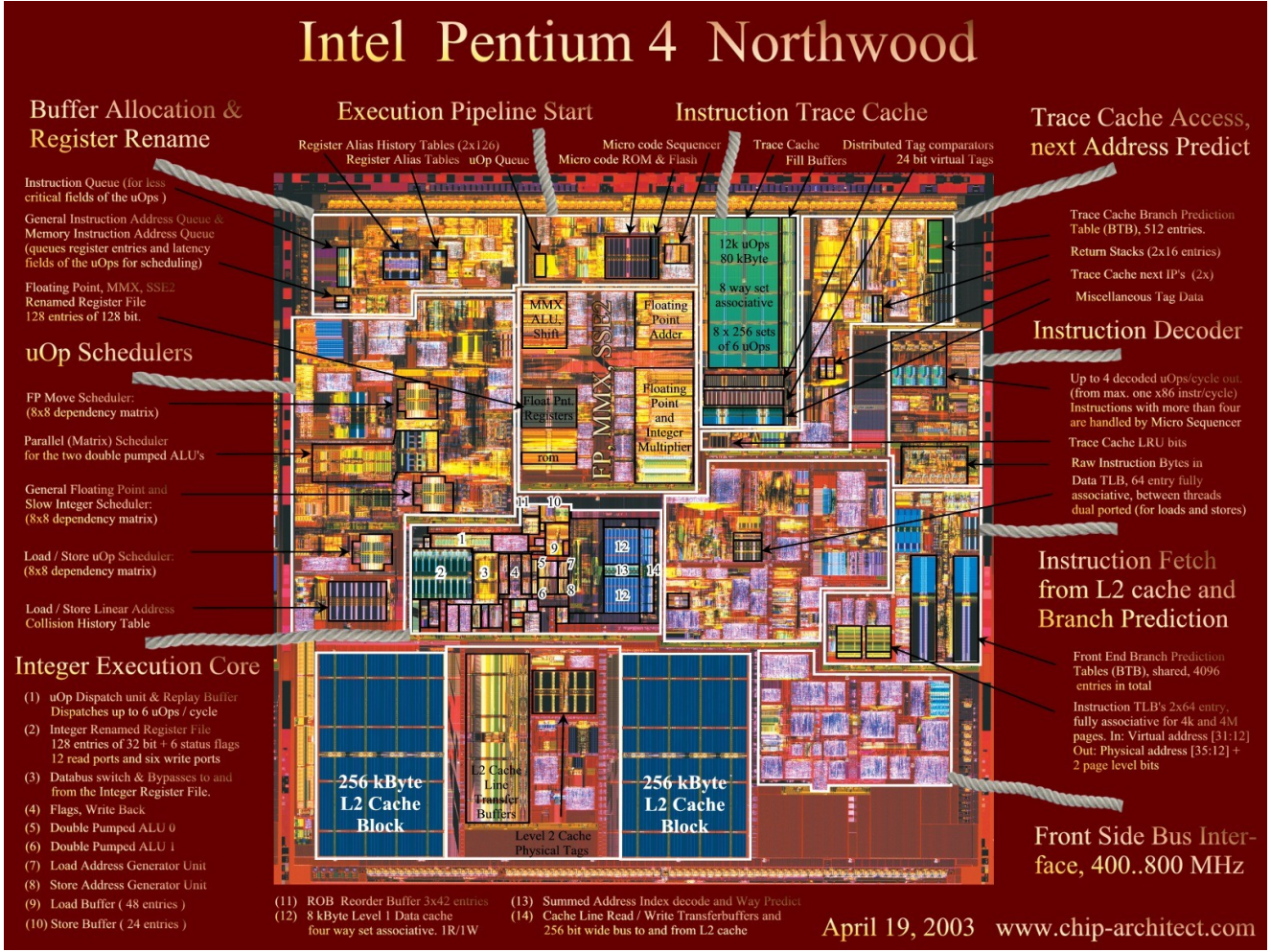
- Multicore processors are ubiquitous
- **Needed to fully utilise all cores of a processor**
- Only way for chip manufacturers to improve performance
- Some processors are practically useless without it
- Problem is too large to fit one machine
- Can even be useful on single core machines

Why parallel computing?

- Multicore processors are ubiquitous
- Needed to fully utilise all cores of a processor
- **Only way for chip manufacturers to improve performance**
- Some processors are practically useless without it
- Problem is too large to fit one machine
- Can even be useful on single core machines



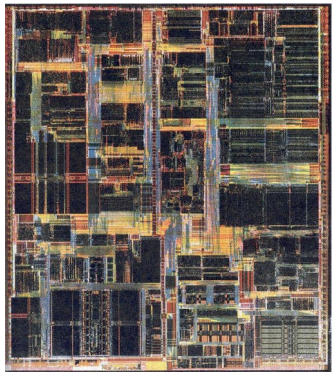

14nm
10nm
7nm
5nm


Intel Pentium 4 Northwood
Buffer Allocation & Register Rename
Execution Pipeline Start
Instruction Trace Cache
Trace Cache Access, next Address Predict
Register Alias History Tables (2x126)
Register Alias Tables
uOp Queue
Micro code Sequencer
Micro code ROM & Flash
Trace Cache Fill Buffers
Distributed Tag comparators 24 bit virtual Tags
Instruction Queue (for less critical fields of the uOps)
General Instruction Address Queue & Memory Instruction Address Queue (queues register entries and latency fields of the uOps for scheduling)
Floating Point, MMX, SSE2 Renamed Register File 128 entries of 128 bit.
12k uOps 80 kByte
8 way set associative
8 x 256 sets of 6 uOps
Trace Cache Branch Prediction Table (BTB), 512 entries.
Return Stacks (2x16 entries)
Trace Cache next IP's (2x)
Miscellaneous Tag Data
MMX ALU, Shift
Floating Point Adder
FP, MMX, SSE2
Float Pnt. Registers
rom
Floating Point and Integer Multiplier
uOp Schedulers
Instruction Decoder
FP Move Scheduler: (8x8 dependency matrix)
Parallel (Matrix) Scheduler for the two double pumped ALU's
General Floating Point and Slow Integer Scheduler: (8x8 dependency matrix)
Load / Store uOp Scheduler: (8x8 dependency matrix)
Load / Store Linear Address Collision History Table
Up to 4 decoded uOps/cycle out. (from max. one x86 instr/cycle) Instructions with more than four are handled by Micro Sequencer
Trace Cache LRU bits
Raw Instruction Bytes in
Data TLB, 64 entry fully associative, between threads dual ported (for loads and stores)
Instruction Fetch from L2 cache and Branch Prediction
Integer Execution Core
(1) uOp Dispatch unit & Replay Buffer Dispatches up to 6 uOps / cycle
(2) Integer Renamed Register File 128 entries of 32 bit + 6 status flags 12 read ports and six write ports
(3) Databus switch & Bypasses to and from the Integer Register File.
(4) Flags, Write Back
(5) Double Pumped ALU 0
(6) Double Pumped ALU 1
(7) Load Address Generator Unit
(8) Store Address Generator Unit
(9) Load Buffer (48 entries)
(10) Store Buffer (24 entries)
256 kByte L2 Cache Block
L2 Cache Line Transfer Buffers
256 kByte L2 Cache Block
Level 2 Cache Physical Tags
Front End Branch Prediction Tables (BTB), shared, 4096 entries in total
Instruction TLB's 2x64 entry, fully associative for 4k and 4M pages. In: Virtual address [31:12] Out: Physical address [35:12] + 2 page level bits
Front Side Bus Interface, 400..800 MHz
(11) ROB Reorder Buffer 3x42 entries
(12) 8 kByte Level 1 Data cache four way set associative. 1R/1W
(13) Summed Address Index decode and Way Predict
(14) Cache Line Read / Write Transferbuffers and 256 bit wide bus to and from L2 cache
April 19, 2003
www.chip-architect.com

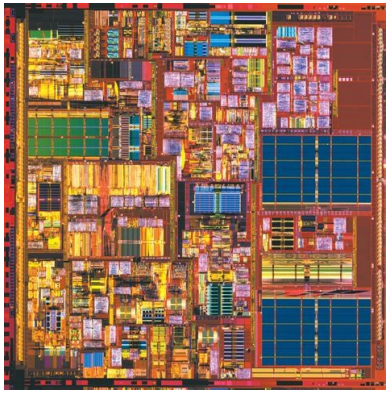
- Improving single core performance linearly requires an exponential number of transistors

- At some point it becomes worth it to spend those transistors on multiple independent cores instead

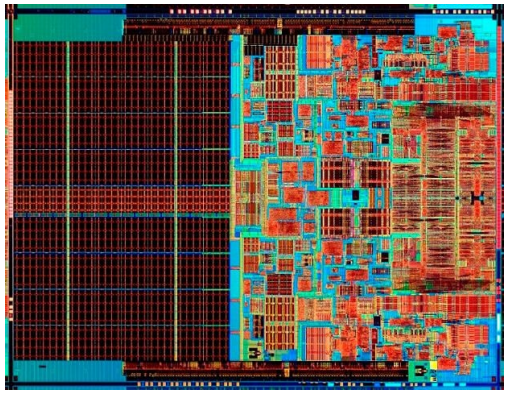
Pentium III



Pentium 4



Pentium D

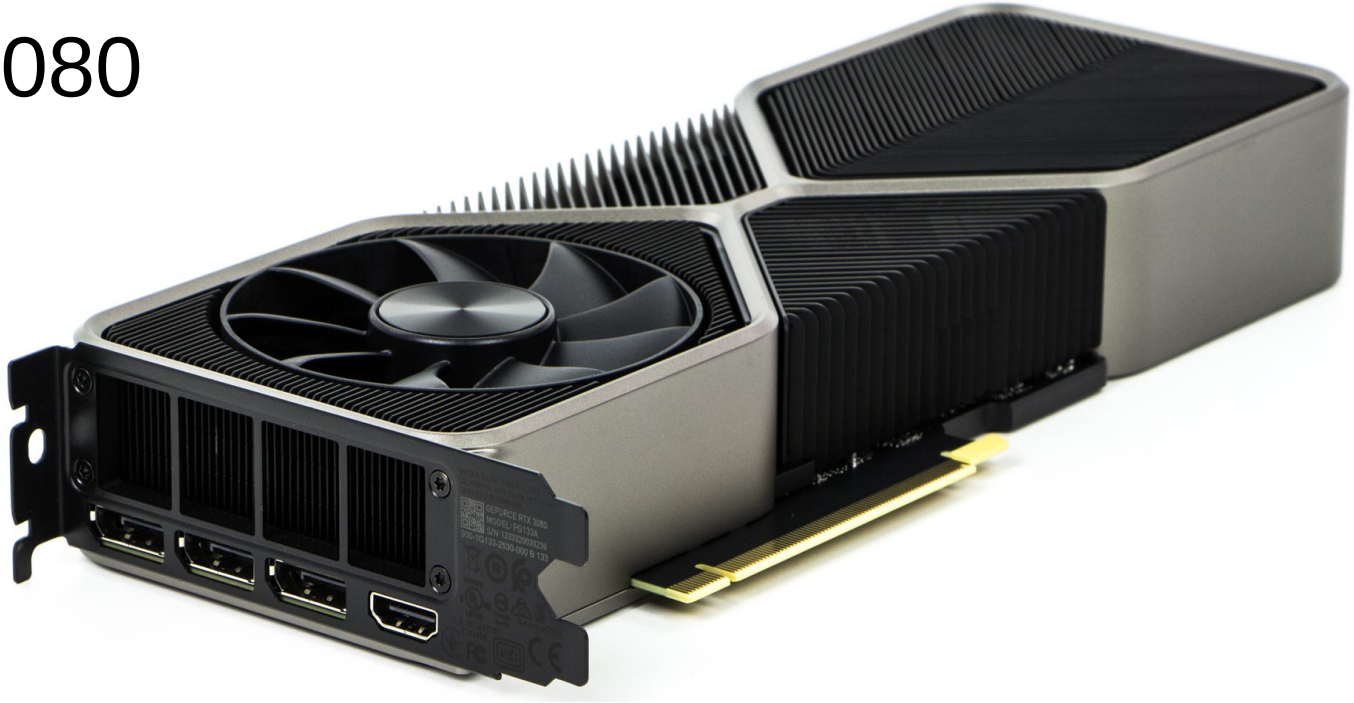


Core 2 Duo

Why parallel computing?

- Multicore processors are ubiquitous
- Needed to fully utilise all cores of a processor
- Only way for chip manufacturers to improve performance
- **Some processors are practically useless without it**
- Problem is too large to fit one machine
- Can even be useful on single core machines

RTX 3080



Power consumption (load): 320W

Theoretical performance: 29,770,000,000 Flops

Theoretical performance: 29,770,000,000 Flops

Theoretical performance: 29,770,000,000 Flops

If you'd do one calculation per second,
the same number of calculations would take:

944 years

A more realistic workload..

DOOM

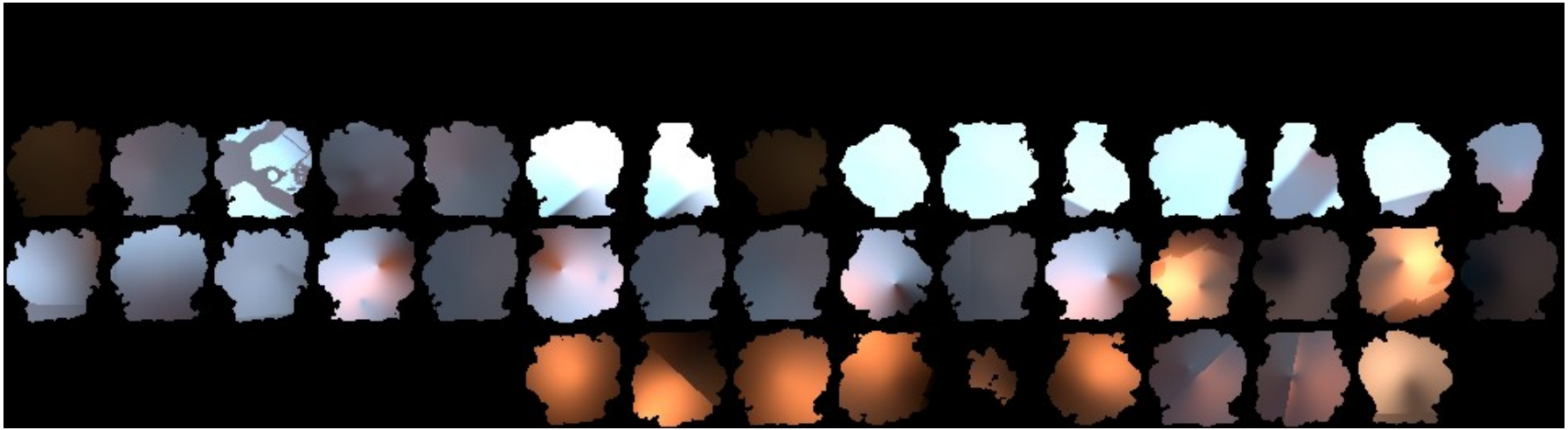
http://www.adriancourreges.com/blog/2016/09/09/doom-2016-graphics-study/

http://www.adriancourreges.com/blog/2016/09/09/doom-2016-graphics-study/

http://www.adriancourreges.com/blog/2016/09/09/doom-2016-graphics-study/

http://www.adriancourreges.com/blog/2016/09/09/doom-2016-graphics-study/

UNION



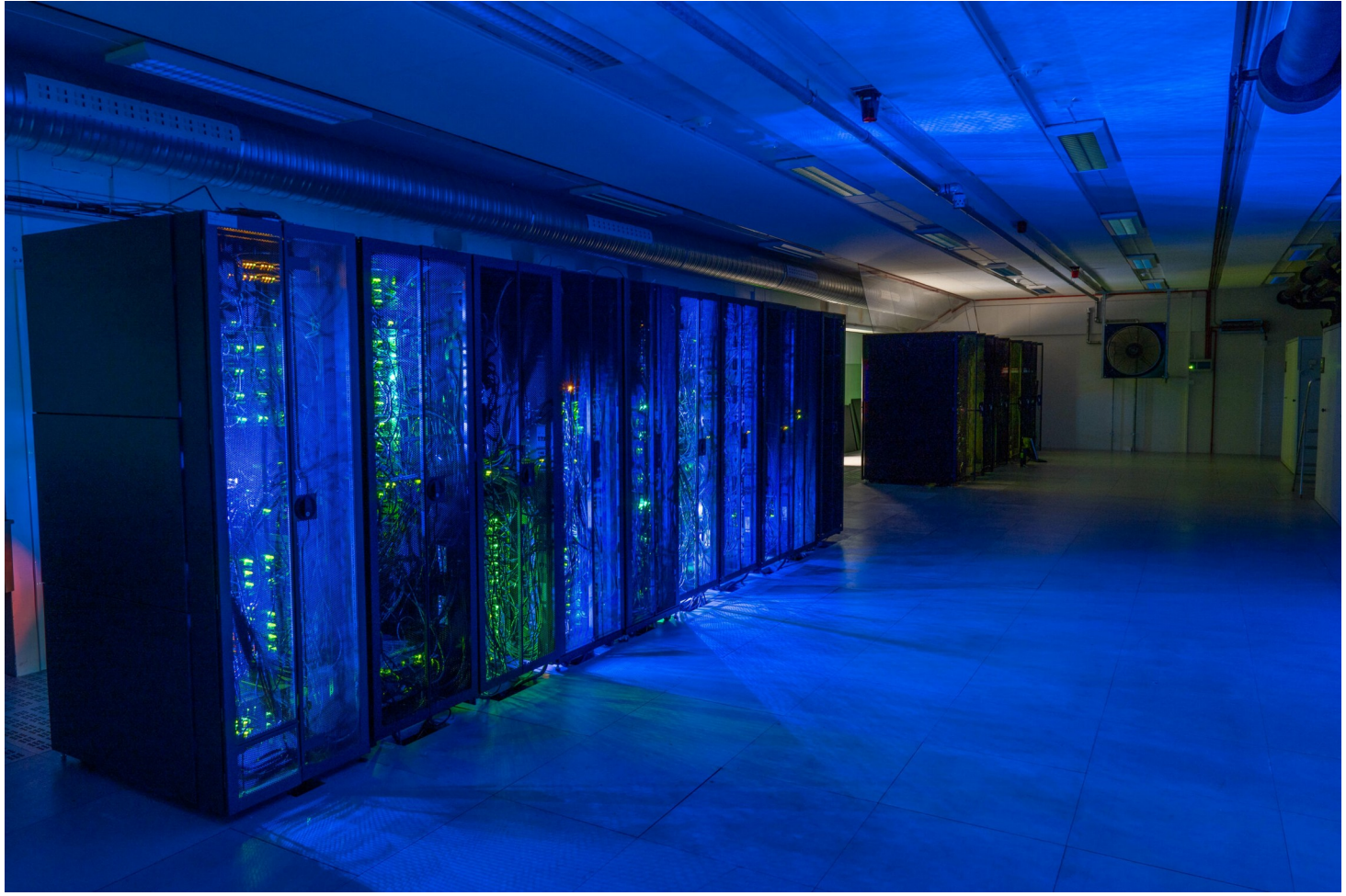
90
9

At a resolution of 2560x1440 pixels,
the GPU does all of that 200 times per second.

(and even gets to take short breaks)

Why parallel computing?

- Multicore processors are ubiquitous
- Needed to fully utilise all cores of a processor
- Only way for chip manufacturers to improve performance
- Some processors are practically useless without it
- **Problem is too large to fit one machine**
- Can even be useful on single core machines



Why parallel computing?

- Multicore processors are ubiquitous
- Needed to fully utilise all cores of a processor
- Only way for chip manufacturers to improve performance
- Some processors are practically useless without it
- Problem is too large to fit one machine
- **Can even be useful on single core machines**

Find the Ground
WAITING FOR OCCLUSION

But why an entire course?

- Communication overhead
- Race conditions

What can happen?

- https://en.m.wikipedia.org/wiki/Therac-25

http://spritesmods.com/?art=hddhack&page=2
https://www.briandorey.com/docs/2020-01-15-sony-55xe9005-teardown/pcb-main.jpg
https://www.rcgeeks.co.uk/blogs/news/dji-mavic-mini-teardown-whats-inside
https://www.komplett.no/img/p/800/1219388.jpg
https://www.ebay.com/itm/402730890541
https://www.ifixit.com/Teardown/GoPro+HERO11+Black+Mini+Teardown/155069
https://news.satnews.com/2020/08/04/xiphos-reveals-their-new-space-processor-board-for-sdr-applications/
https://www.hpc.ntnu.no/idun/
https://cdn.benchmark.pl/uploads/backend_img/c/newsy/2020-09/PM/nvidia-rtx-3080_05.jpg
https://pbs.twimg.com/media/ELcxWo0U8AAAsiR.jpg
http://www.chip-architect.org/news/Northwood_130nm_die_text_1600x1200.jpg
https://www.zeiss.com/spectroscopy/applications-industries/oem-applications/semiconductor.html
https://wp.technologyreview.com/wp-content/uploads/2018/08/googlecbf009-11.jpg
https://tpucdn.com/cpu-specs/images/chips/2817-die-shot.jpg
https://www.techrepublic.com/wp-content/uploads/2011/11/22inteldieshot1997.jpg
http://brainstones.narod.ru/collection/intel/intel_pentium_d_925_sl9ka_wo_lid.jpg