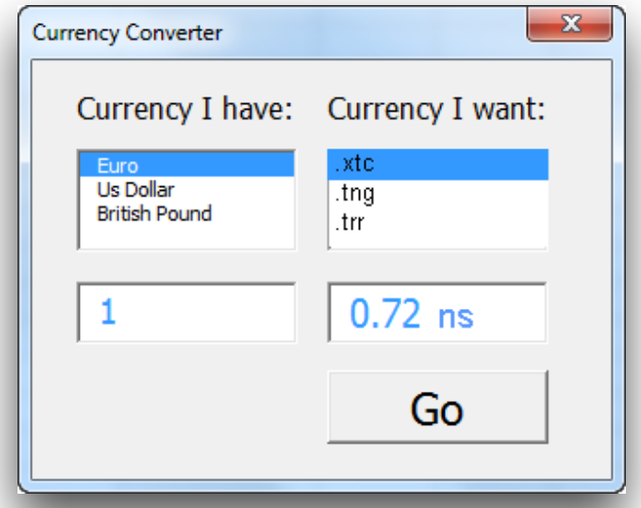
Outperforming Professional HPC With Consumer GPUs and Hardware Re-Use

Carsten Kutzner, Szilárd Páll, Martin Fechner, Ansgar Esztermann, Bert de Groot, Helmut Grubmüller

Motivation



- **fixed budget** for compute hardware
- How to optimally make use of that?
  - We run mostly **GROMACS** MD,
    ➔ tailor nodes for GROMACS
  - queue is always full ➔ optimise for **throughput / single-node performance**
  - (scaling ➔ HPC centres)
  - how to **produce a maximum amount of MD trajectory per invested €** over 3-5 years?

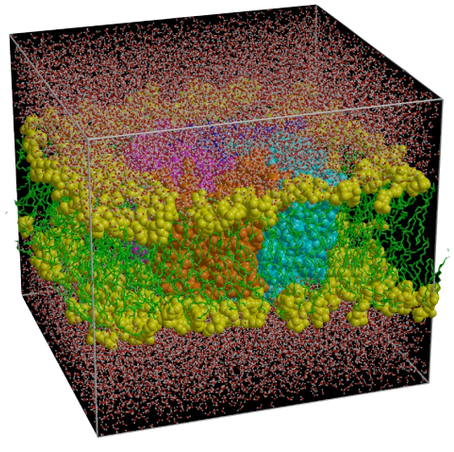
Approach

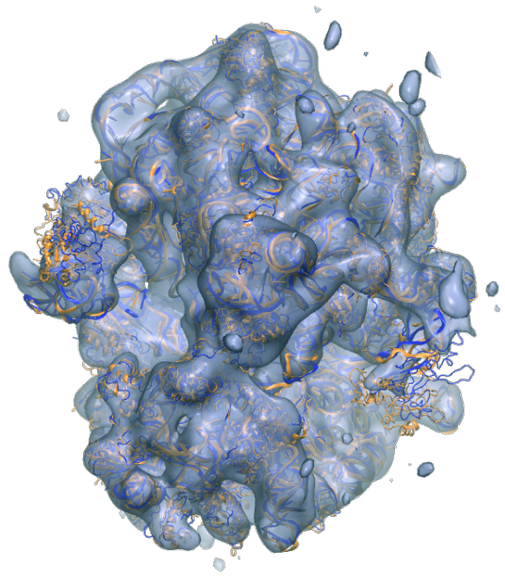
- from ~10 CPU types + ~10 GPU models we **assemble** and **benchmark** various compute nodes
  - CPU nodes
  - GPU nodes with 1, 2, 3, and 4 GPUs
  - **consumer** and **professional** GPUs
- determine **performance-to-price ratio**

- on multi-GPU nodes, benchmarks use 1 simulation per GPU,
  - reported node performance (ns/d) is sum of the performances of the individual simulations (**“aggregate” performance**)

- benchmark MD systems:

80k atom MEM benchmark
channel in membrane + water + ions, PME, 2 fs time step





2M atoms RIB benchmark
ribosome in solution, PME, 4 fs time step

consumer GPUs (GeForce)
- GTX 980
- GTX 1070
- GTX 1070Ti
- GTX 1080
- GTX 1080Ti
- RTX 2070
- RTX 2080
- RTX 2080Ti
- …

professional GPUs (Tesla)
- Quadro P6000
- Tesla V100
- …

CPUs
- Ryzen (16 core)
- Epyc (24 core)
- Core i7 (4 core)
- Xeon (4, 6, 8, 10, and 20 core)
- …

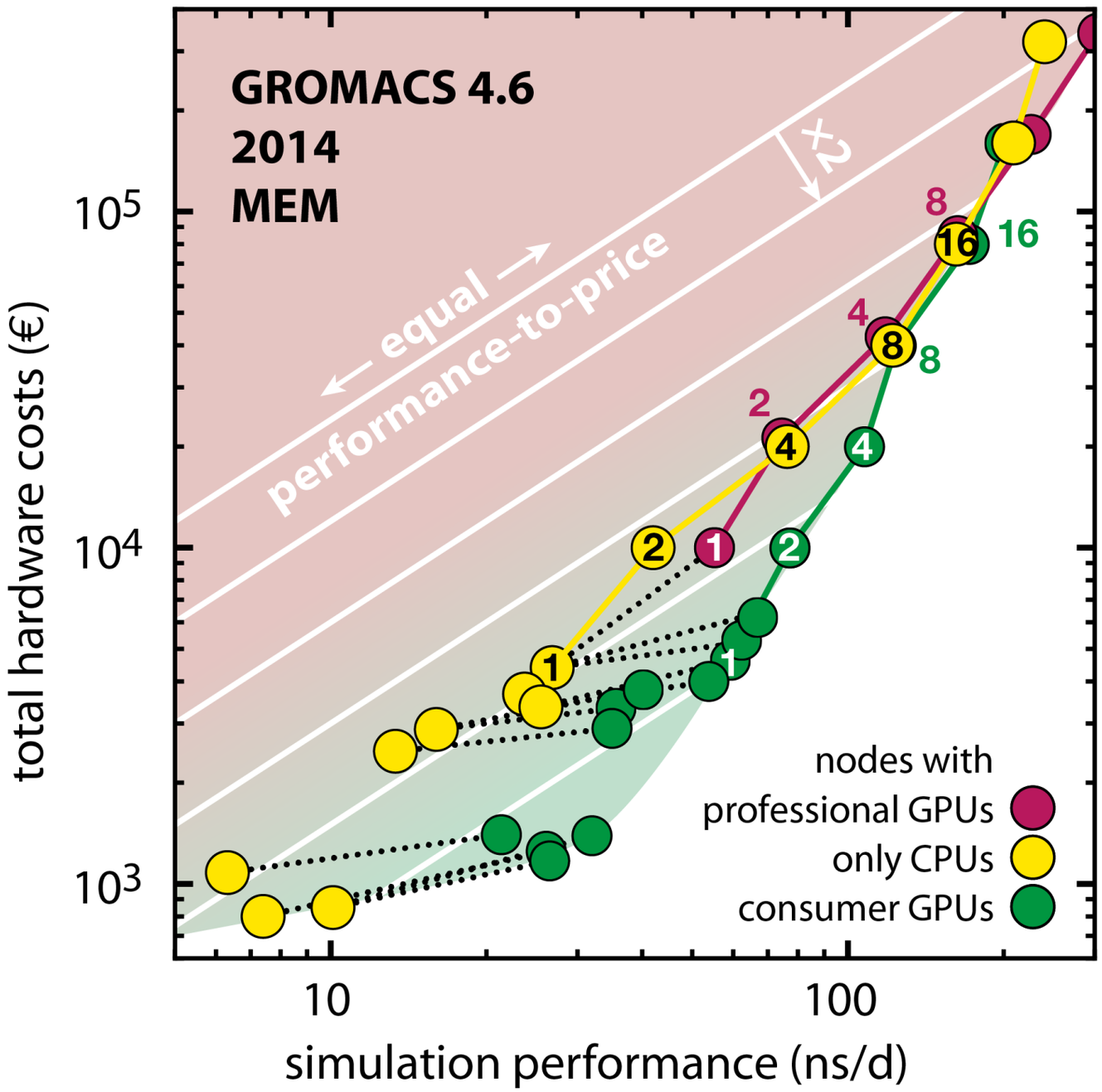
2014: First Comprehensive Hardware Evaluation

- Main 2014 result:

  ● nodes with GeForce consumer GPUs

  produce **2–3x** as much MD trajectory per invested € as

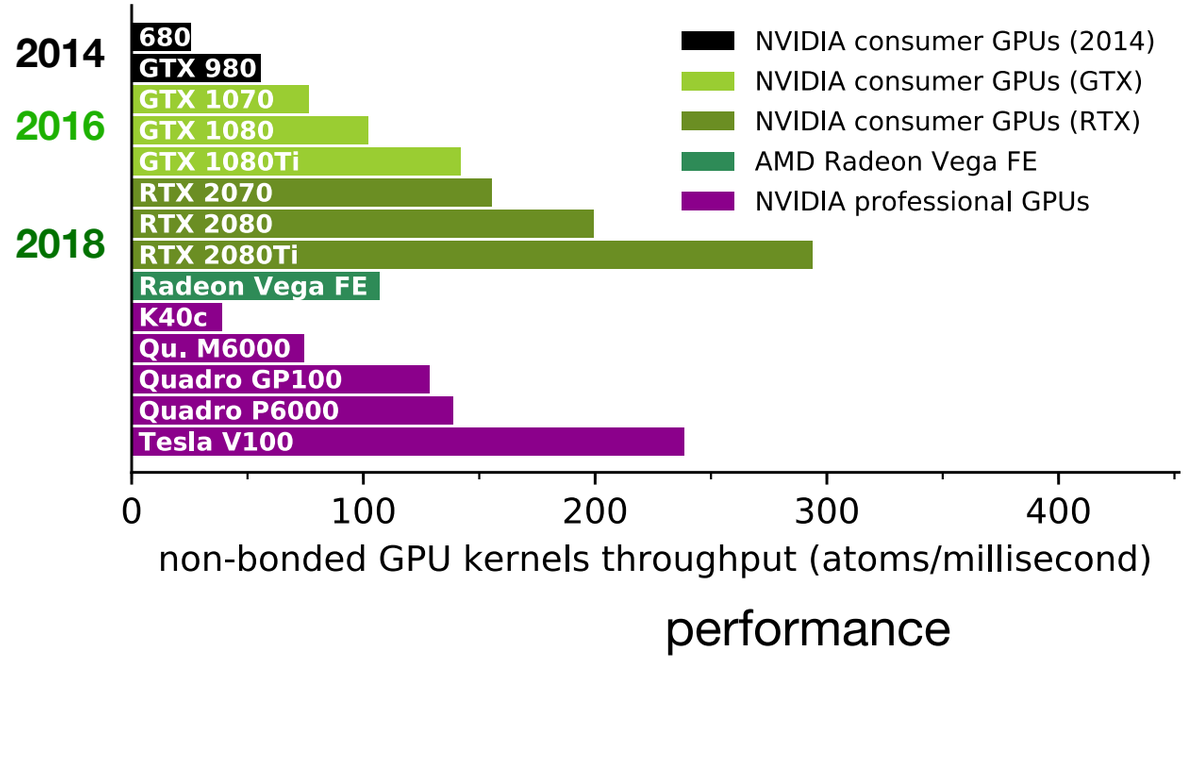
  ● CPU nodes



C Kutzner, S Páll, M Fechner, A Esztermann, BL de Groot, H Grubmüller.
Best bang for your buck: GPU nodes for GROMACS biomolecular simulations.
JCC 36 (26), pp. 1990 - 2008 (2015)

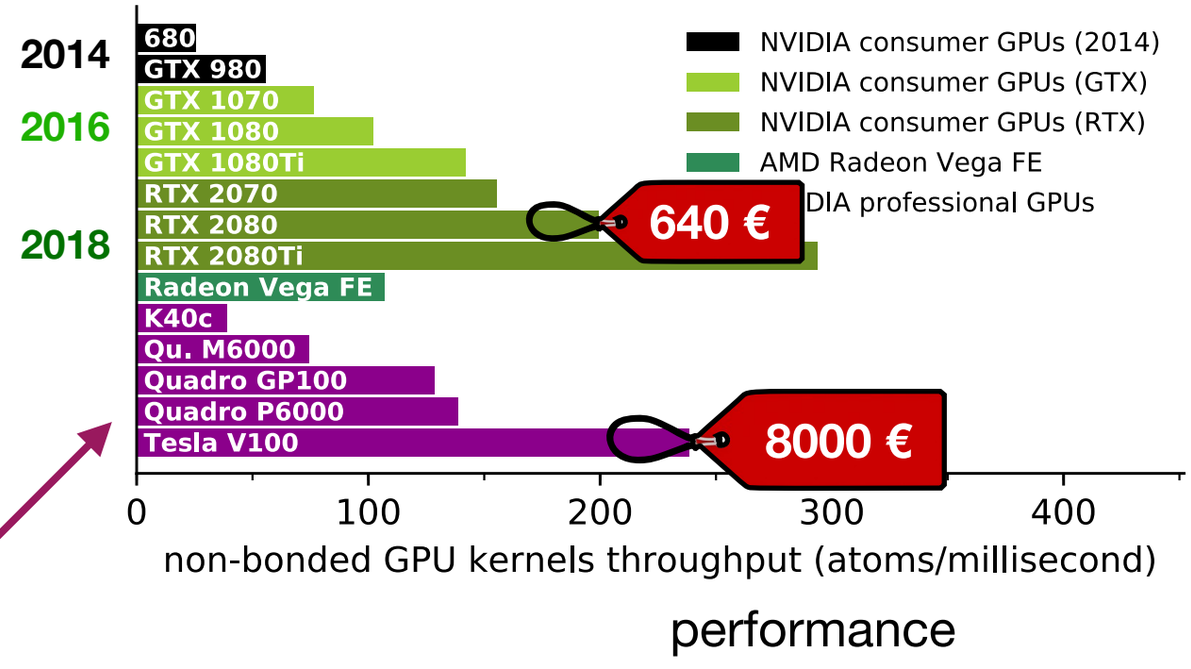
Hardware Developments Since 2014

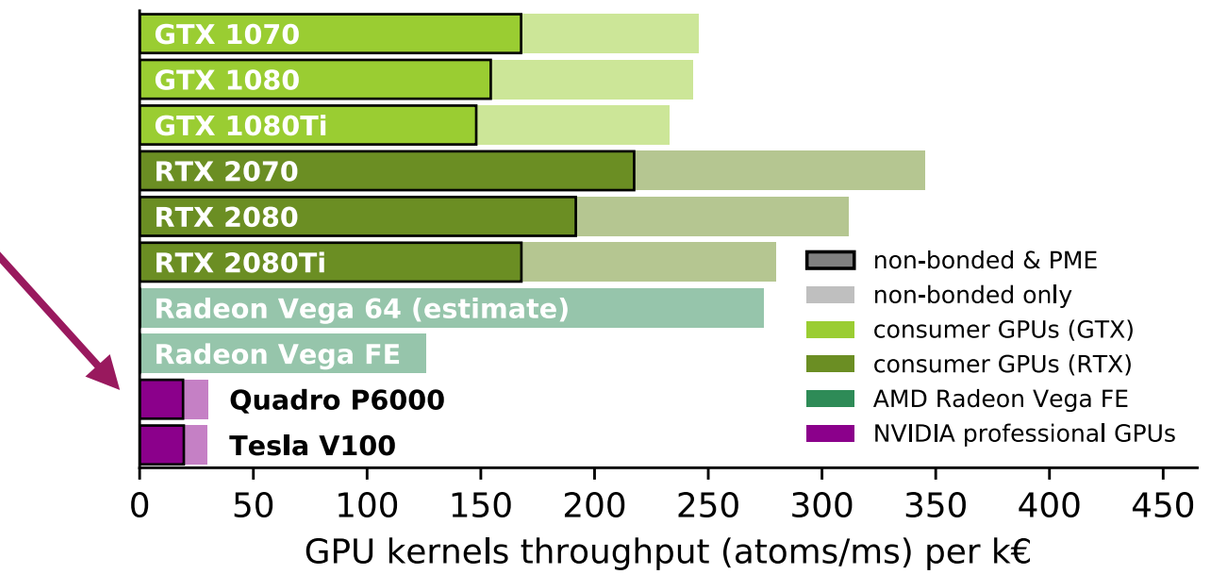
- FLOP-based GPU processing power x3!
- + microarchitectural improvements: up to **6x performance increase** in GPU kernels
- CPU performance: only modest gains



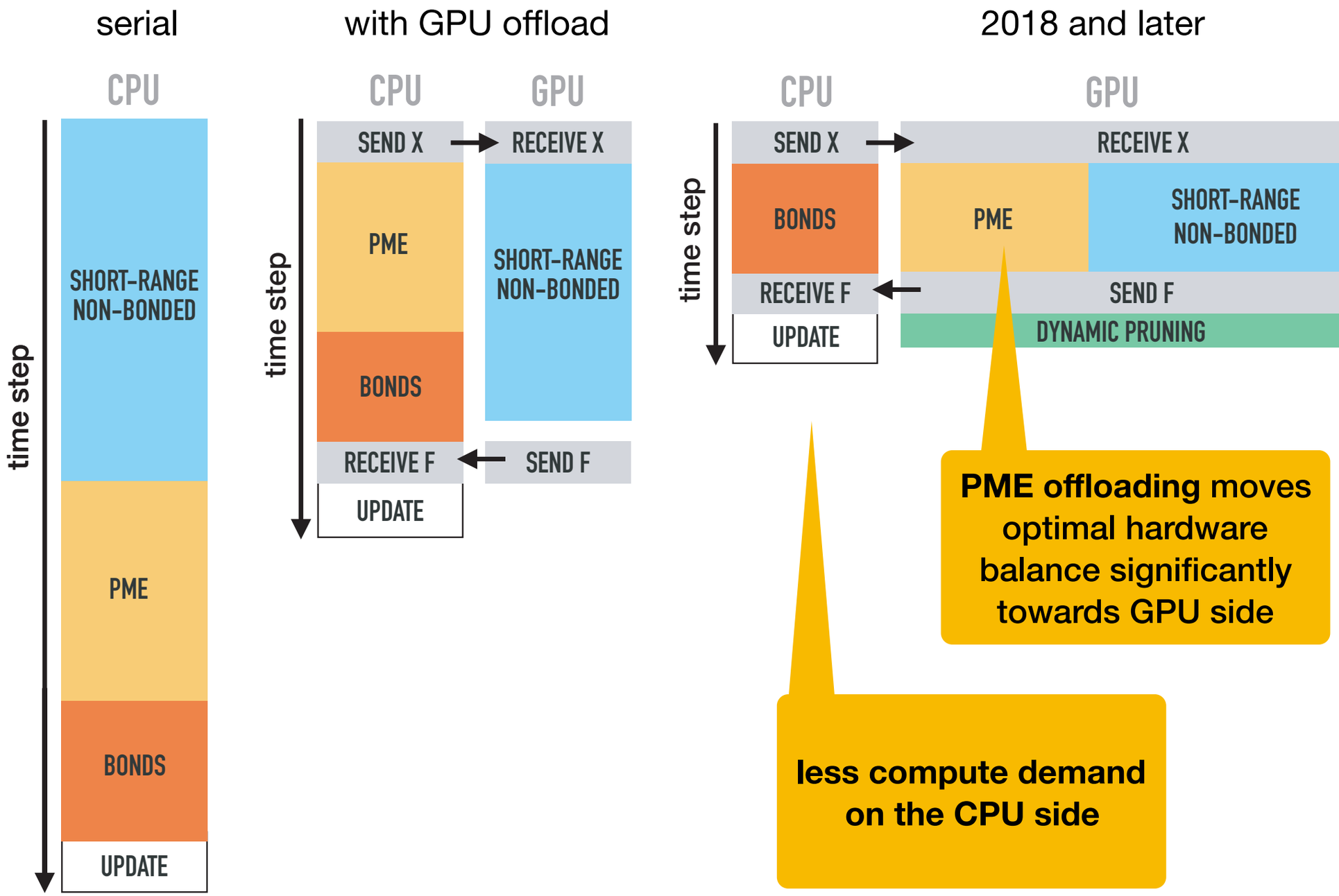
Hardware Developments Since 2014

- FLOP-based GPU processing power x3!
- + microarchitectural improvements: up to **6x performance increase** in GPU kernels
- CPU performance: only modest gains
- **Professional Tesla GPUs** compete with **consumer GPUs** in terms of performance, but are lagging far behind in terms of performance-to-price



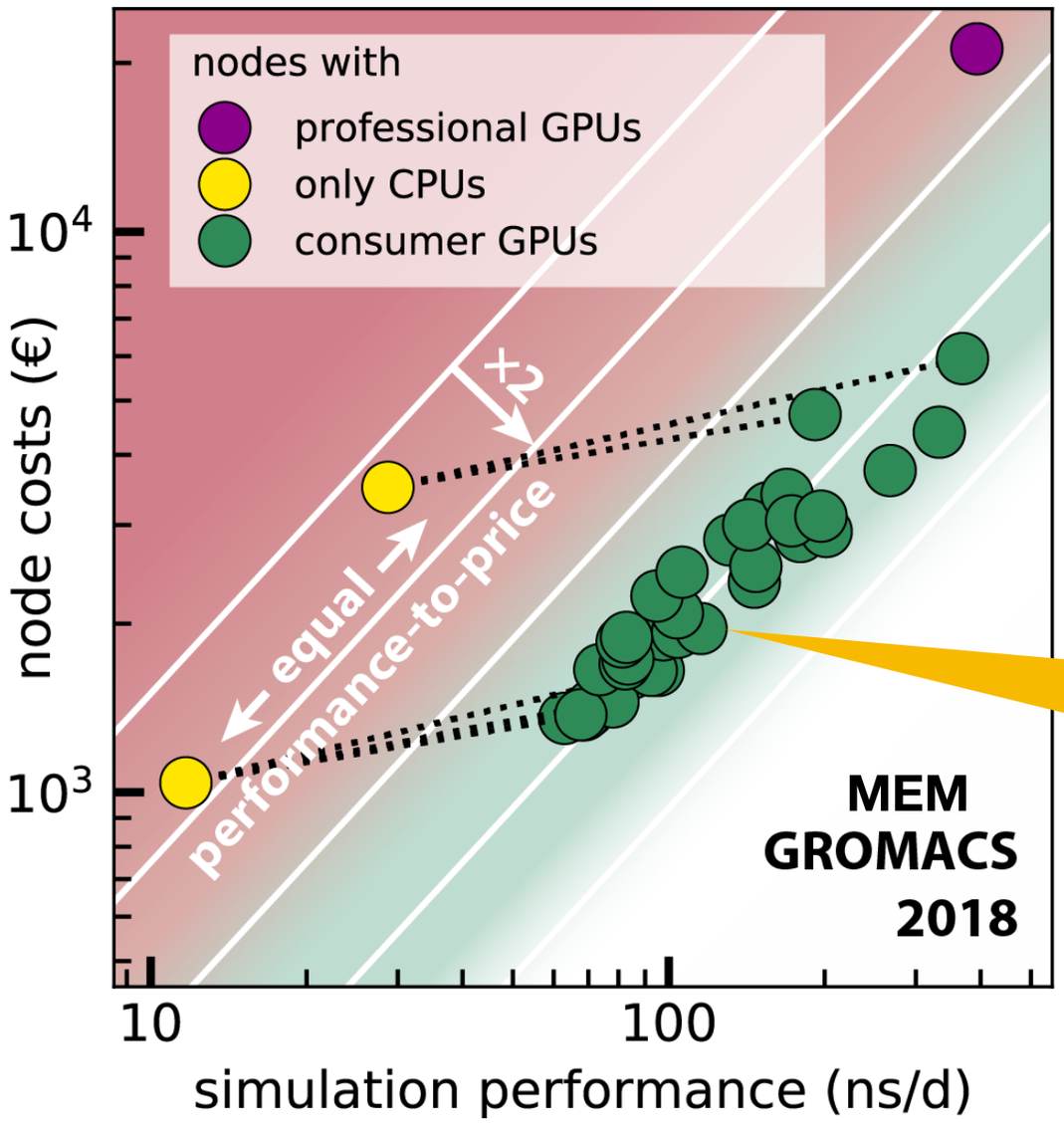


Software Developments: PME Offloading



Páll, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E.: **Tackling exascale software challenges in molecular dynamics simulations with GROMACS.** EASC, Stockholm, Sweden, Revised Selected Papers, pp. 3 – 27 (Eds. Markidis, S.; Laure, E.). Springer, Cham (2015)

The Gap Widens With GROMACS 2018



- Main 2014 result:

  nodes with GeForce consumer GPUs

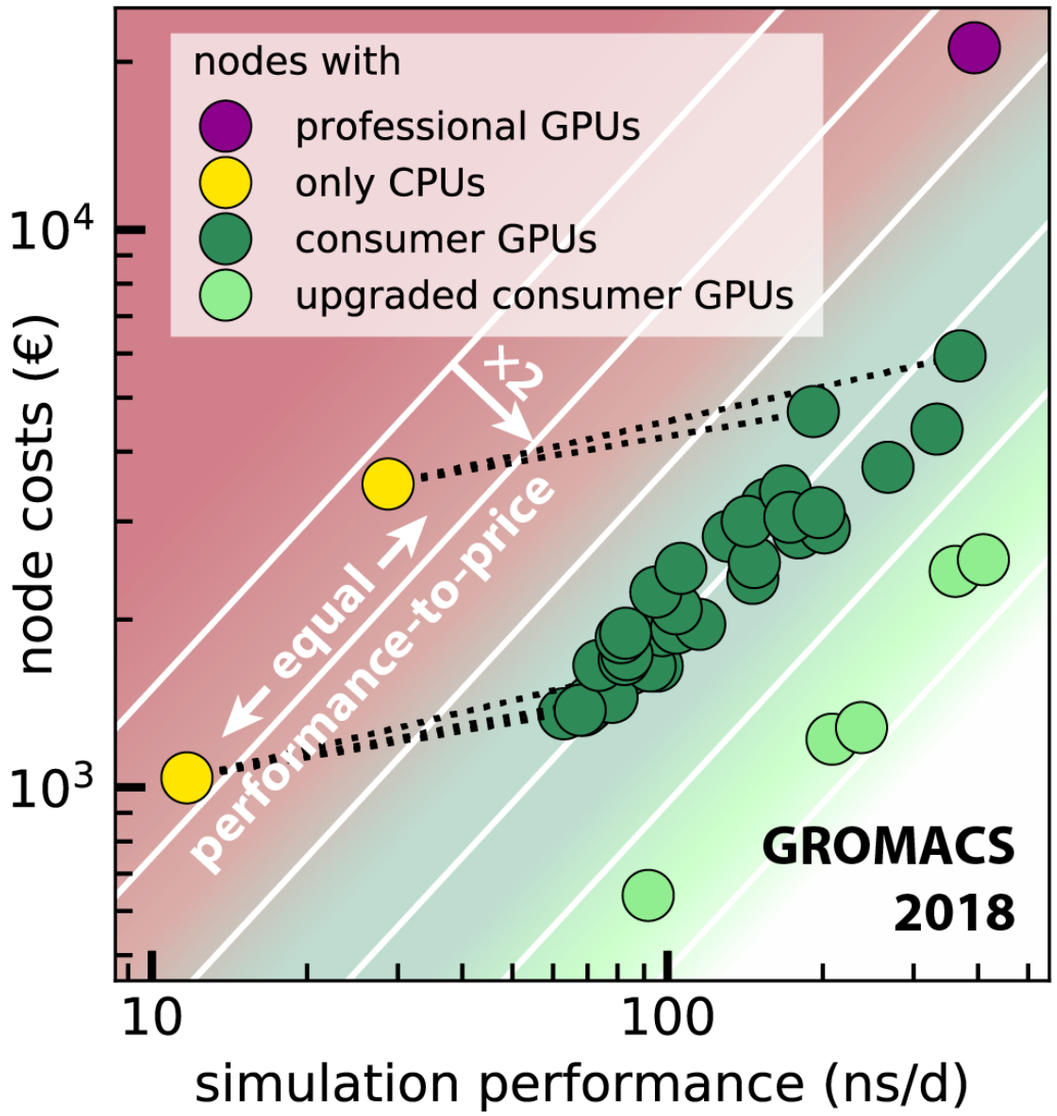
  produce **2–3x** as much MD trajectory per invested € as

  CPU nodes

3–6 x with GROMACS 2018

C Kutzner, S Páll, M Fechner, A Esztermann, BL de Groot, H Grubmüller.
More bang for your buck: Improved use of GPU nodes for GROMACS 2018.
Manuscript under review @ JCC

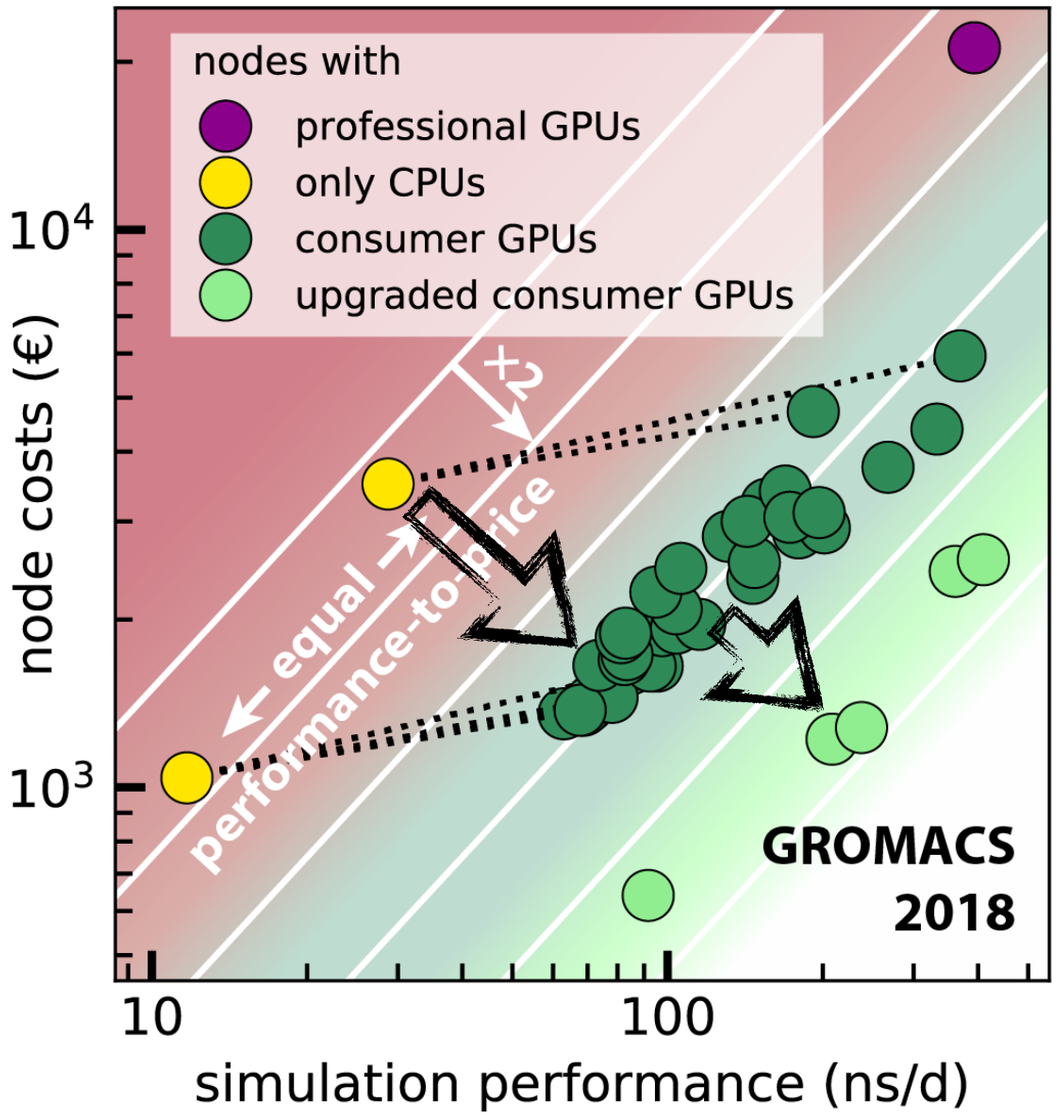
Free Lunch! GPU Upgrades



- shift CPU ➔ GPU **allows to upgrade old nodes with recent GPUs**!

- e.g. E3-1270v2 CPU
  (4 cores @3.5 GHz)
  + GTX 680 (27 ns/d)
  + (●) RTX 2080 (92 ns/d) ➔ 3.4x perf!

Free Lunch! GPU Upgrades



- shift CPU ➔ GPU **allows to upgrade old nodes with recent GPUs**!

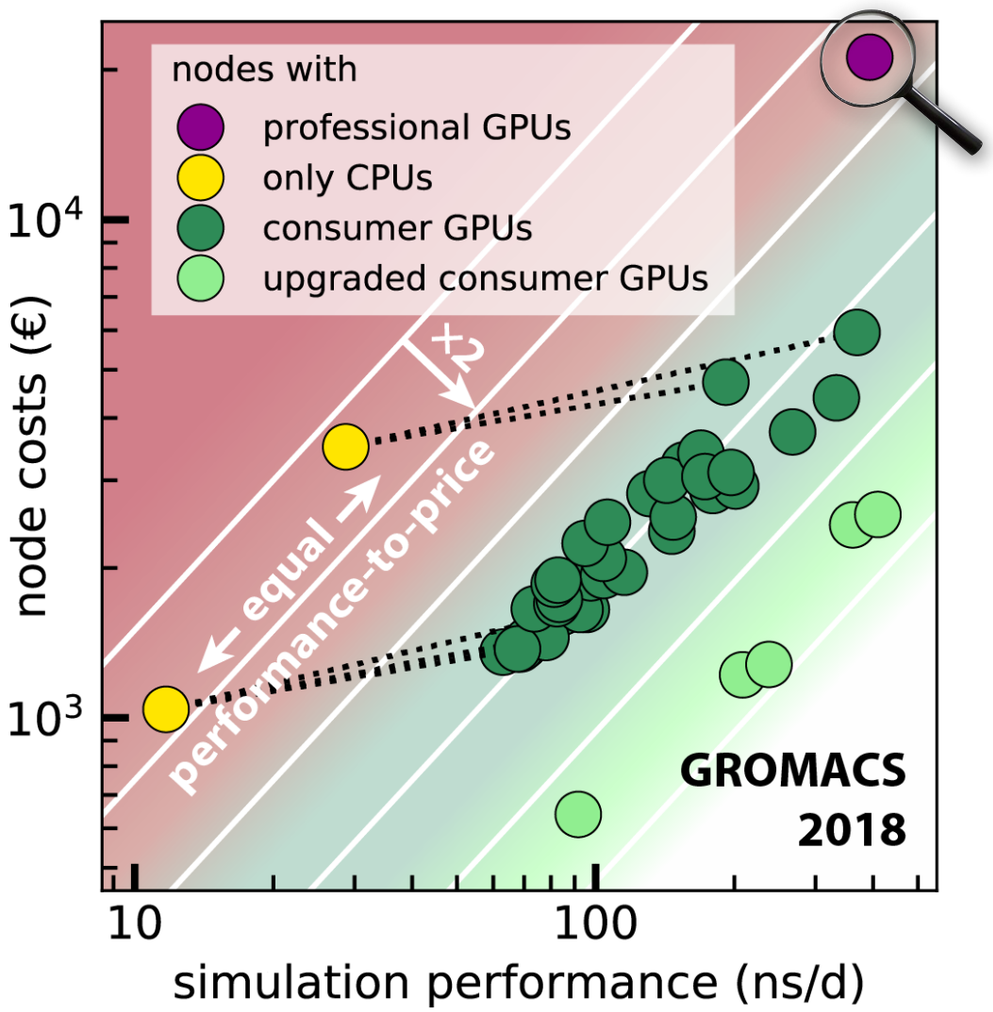
- e.g. E3-1270v2 CPU
  (4 cores @3.5 GHz)
  + GTX 680 (27 ns/d)
  + (●) RTX 2080 (92 ns/d) ➔ 3.4x perf!

- a second leap in performance-to-price

How to Optimally Invest 100 k€ ?

Examples:

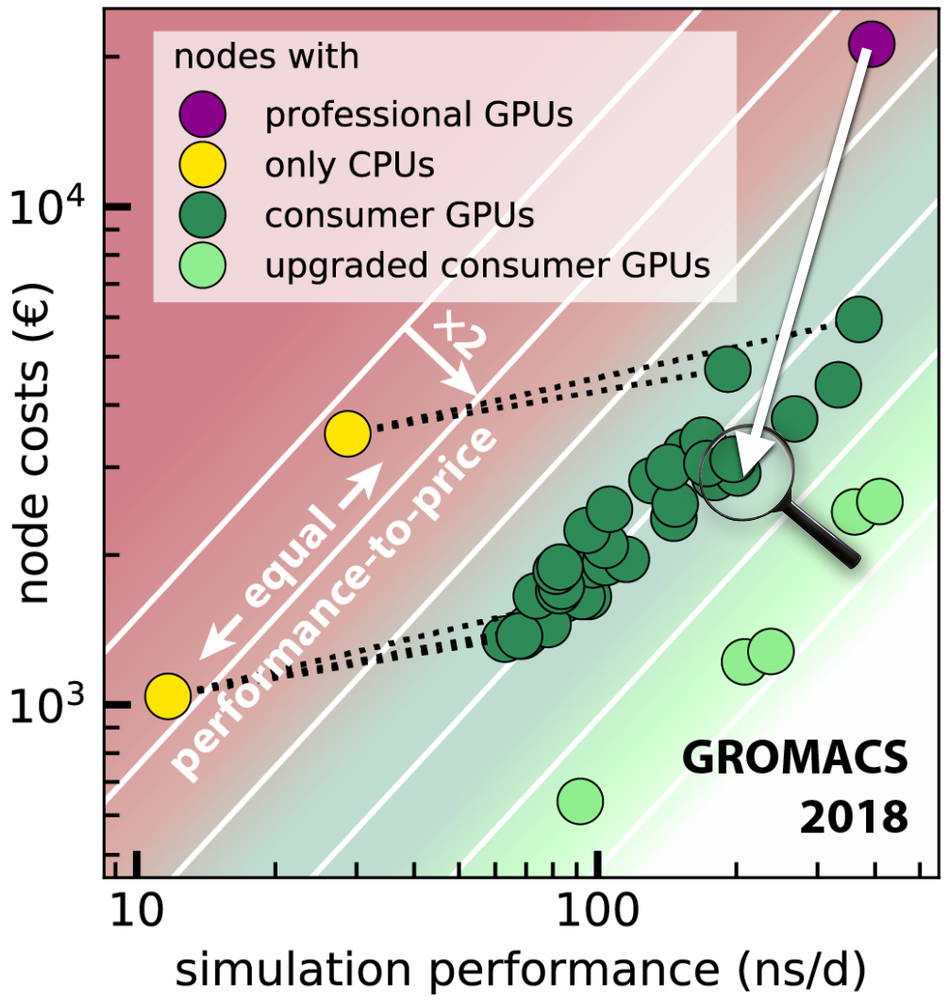
node	CPU	cores	GHz	GPU	perf. (ns/d)	node cost	# nodes	cluster µs/d
new	Gold6148F x2	2 x 20	2.4	V100 x2	393	23200 €	4,3	**1,7**



How to Optimally Invest 100 k€ ?

Examples:

	node	CPU	cores	GHz	GPU	perf. (ns/d)	node cost	# nodes	cluster µs/d
●	new	Gold6148F x2	2 x 20	2.4	V100 x2	393	23200 €	4,3	**1,7**
●	new	E5-2630v4	10	2.2	2080 x2	201	2920 €	34,2	**6,9**



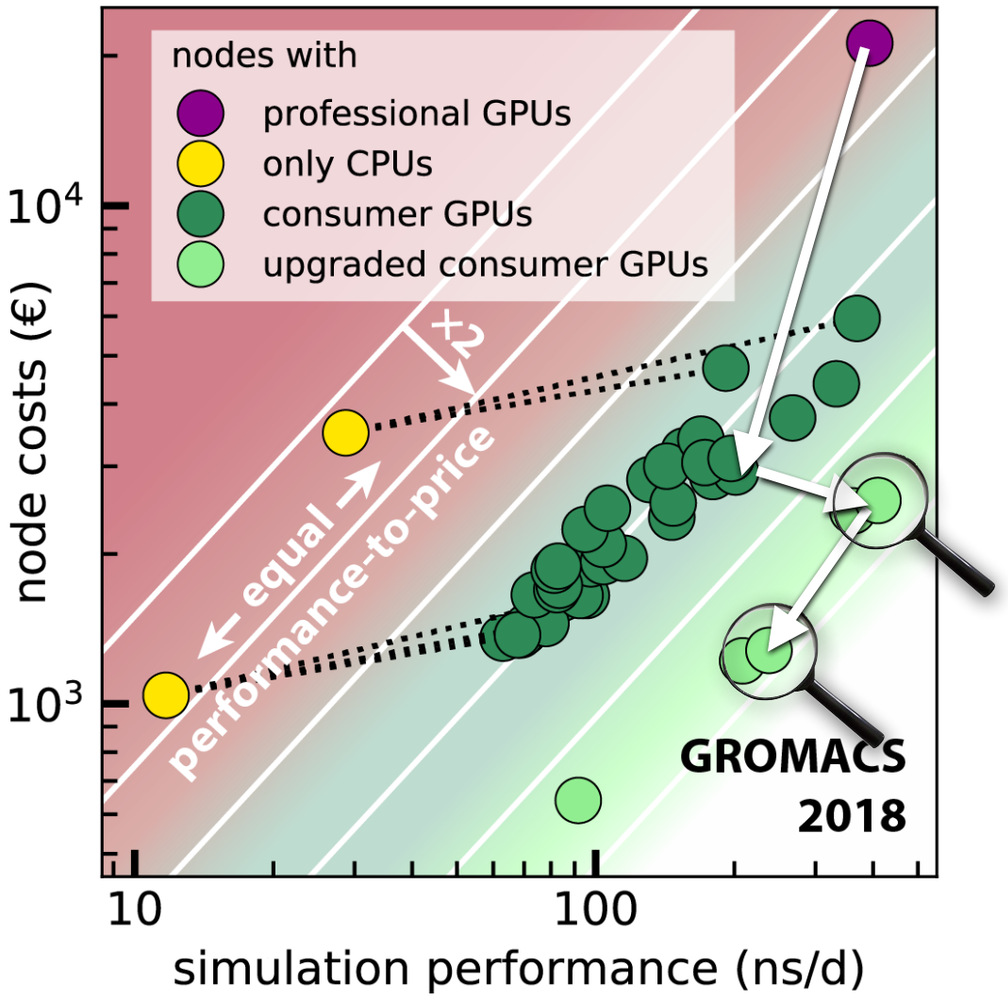
- Consumer GPU (●) nodes produce 4x as much trajectory as general-purpose (●) HPC nodes with Tesla GPUs

How to Optimally Invest 100 k€ ?

Examples:

	node	CPU	cores	GHz	GPU	perf. (ns/d)	node cost	# nodes	cluster µs/d
●	new	Gold6148F x2	2 x 20	2.4	V100 x2	393	23200 €	4,3	**1,7**
●	new	E5-2630v4	10	2.2	2080 x2	201	2920 €	34,2	**6,9**
●	re-used	E5-2680v2 x2	2 x 10	2.8	2080 x4	**410**	2560 €	39,1	**16,0**
●	re-used	E5-2680v2 x2	2 x 10	2.8	2080 x2	238	1280 €	78,1	**18,6**

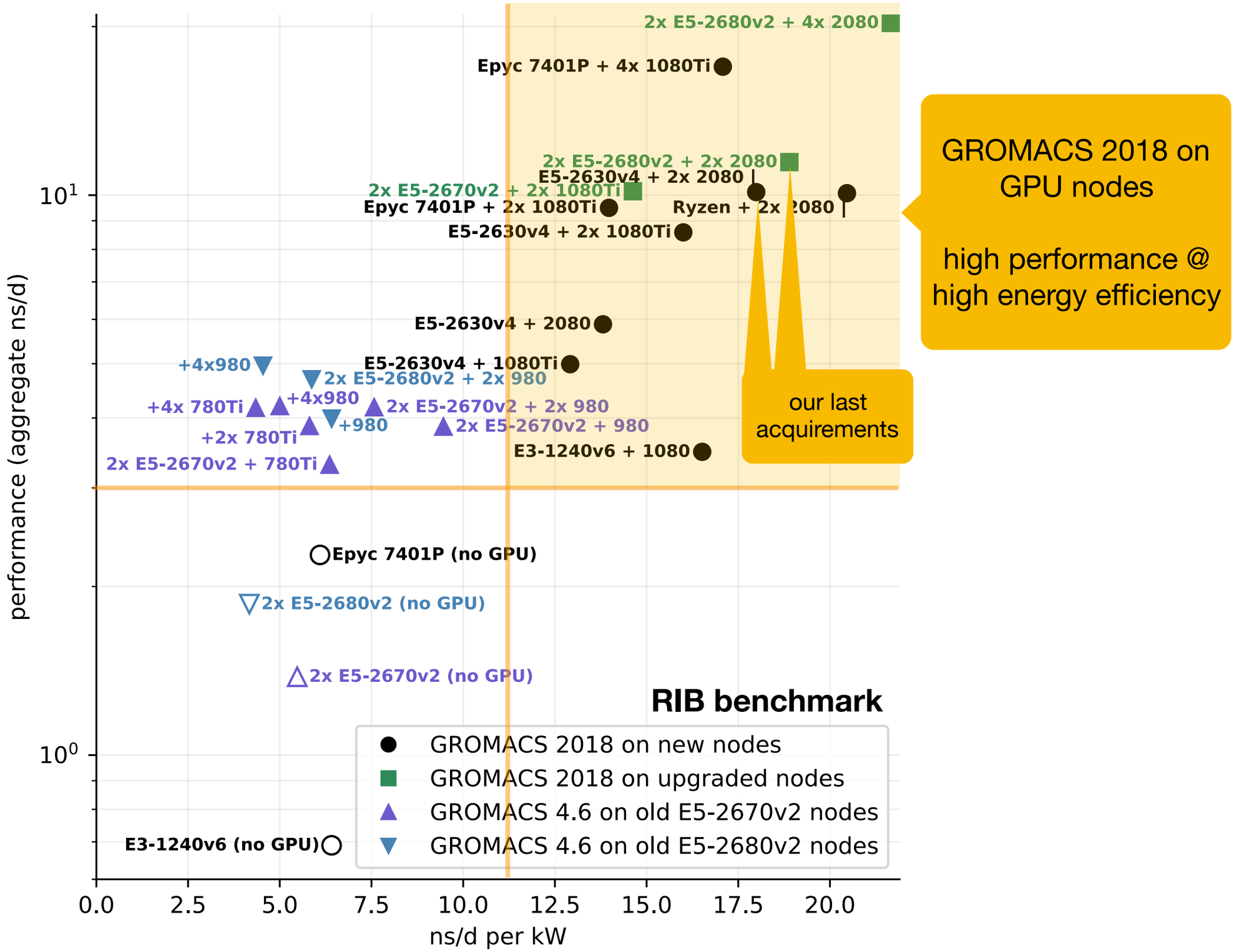
5 year old nodes, decommissioned in 10/2018



- Consumer GPU (●) nodes produce 4x as much trajectory as general-purpose (●) HPC nodes with Tesla GPUs

- Re-used nodes (●), upgraded with consumer GPUs, produce 10x as much trajectory

Performance per Watt



Conclusions for GROMACS 2018

Buying new nodes:

- Consumer GPU nodes have a **3-6x higher performance-to-price ratio** than CPU nodes

Even better: Recycling old nodes! As a result of CPU ➔ GPU work shifting

- **upgrading the GPU** yields large performance increase, whereas
- exchanging the rest of a node (CPU, ..) can be a **waste of money**

Hardware re-use + consumer GPUs allow to tackle large simulation projects
with just a department cluster

The next years …

- GROMACS 2019: bonded interactions ➔ GPU,
  future versions might downgrade the CPU for I/O only
- **Keep our old CPUs+servers,**
  invest in GPUs only

Acknowledgements

The department of Theoretical and Computational Biophysics