lrz

Leibniz Supercomputing Centre

of the Bavarian Academy of Sciences and Humanities

More Compute with Less Energy

Prof. Dr. Dieter Kranzlmüller
Leibniz Supercomputing Centre (LRZ) &
Ludwig-Maximilians-Universität (LMU) Munich

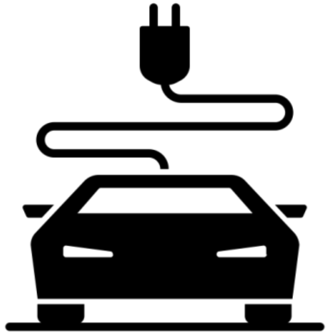
With contributions from
Michael Ott, Axel Auweter,
Josef Weidendorfer,
Nicolay Hammer,
and many others …

How much electrical power do we need for the digital transformation?

- **Smartphone**: iPhone 8
  - Battery Capacity 1821 mAh
  - Per charge (with 0,30 € per kWh) □ approx. 0,0027 € per day
  - With one charge per day □ approx. 0,9855 € per year

Created by Thays Malcher from Noun Project

- Personal computer:
  - Current Intel/AMD multicore processsor with dedicated GPU
  - Per day (with 4 hours running) □ 200 kWh per year
    □ approx. 60 € per year
- Electric vehicle : Tesla
  - Model S 60, 18.1 kWh/100 km □ 5,43 € per 100 km
  - Annual mileage approx. 15.000 km □ 814,50 € per year



Created by Symbolon from Noun Project

Leibniz Supercomputing Centre (LRZ)
of the Bavarian Academy of Sciences and Humanities

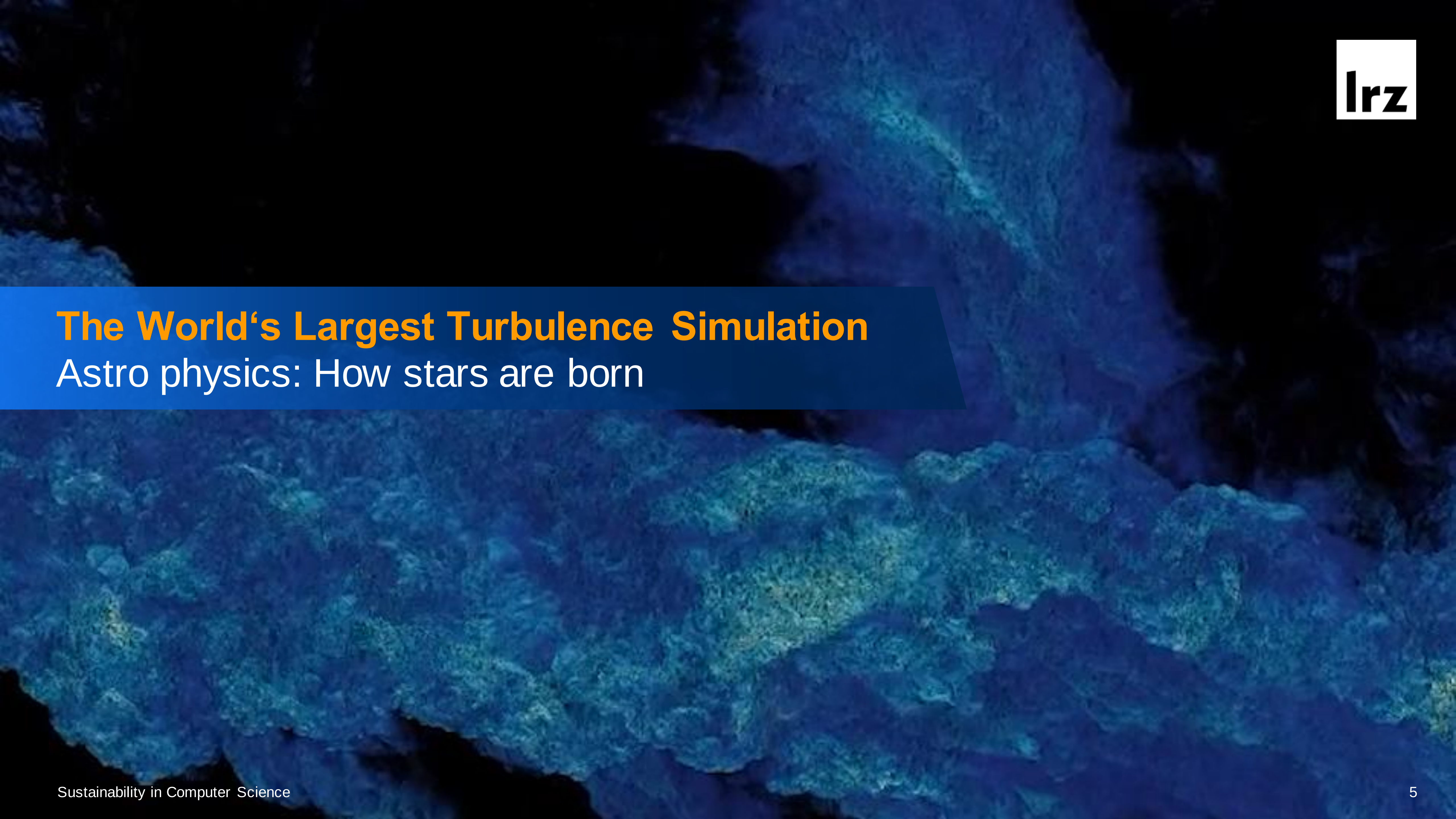
Partner for Digital Transformation of Science & Research

IT-Services for Science since 1962

Munich Universities, Bavaria, Germany, Europe, …

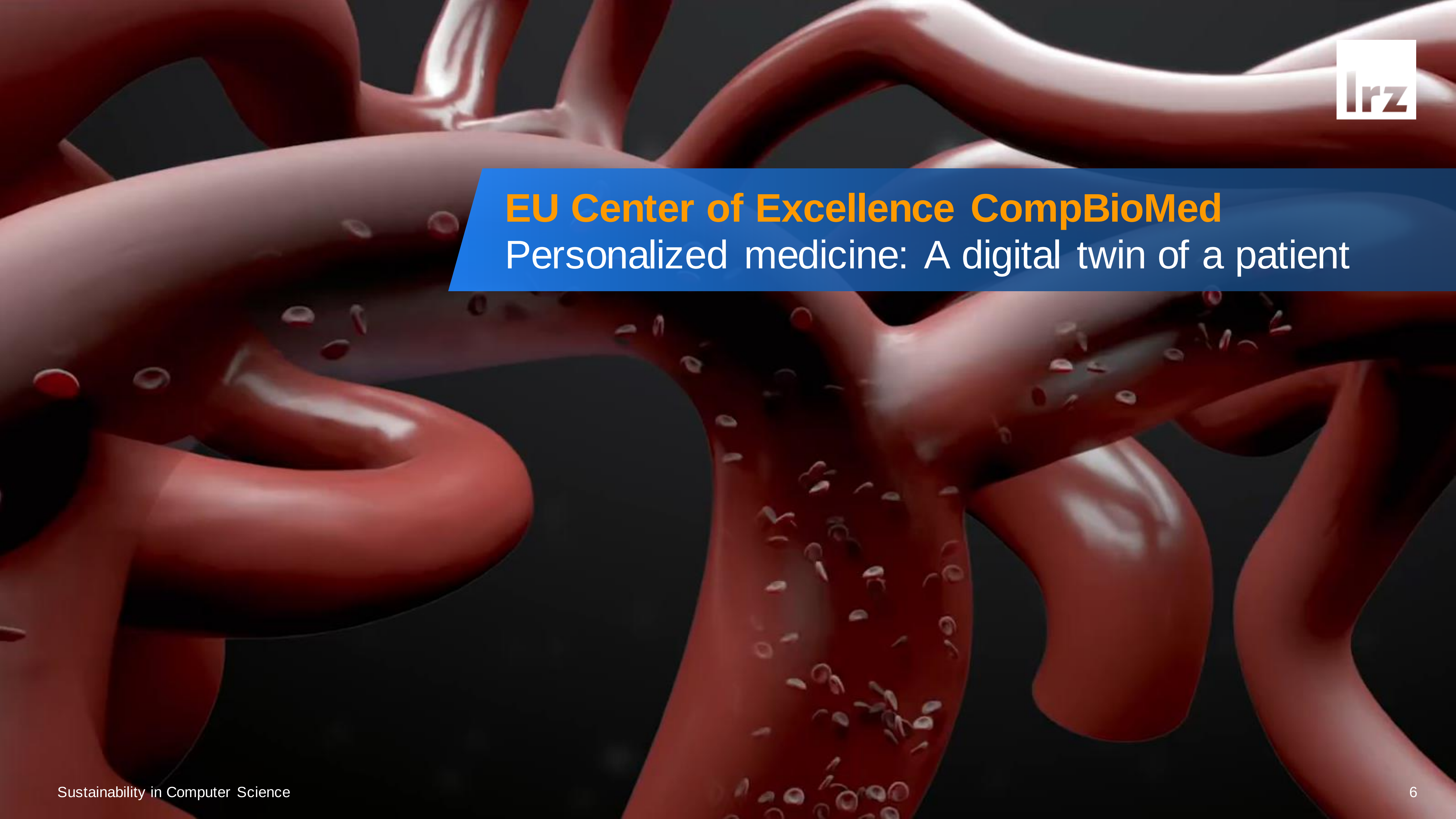
The World‘s Largest Turbulence Simulation

Astro physics: How stars are born

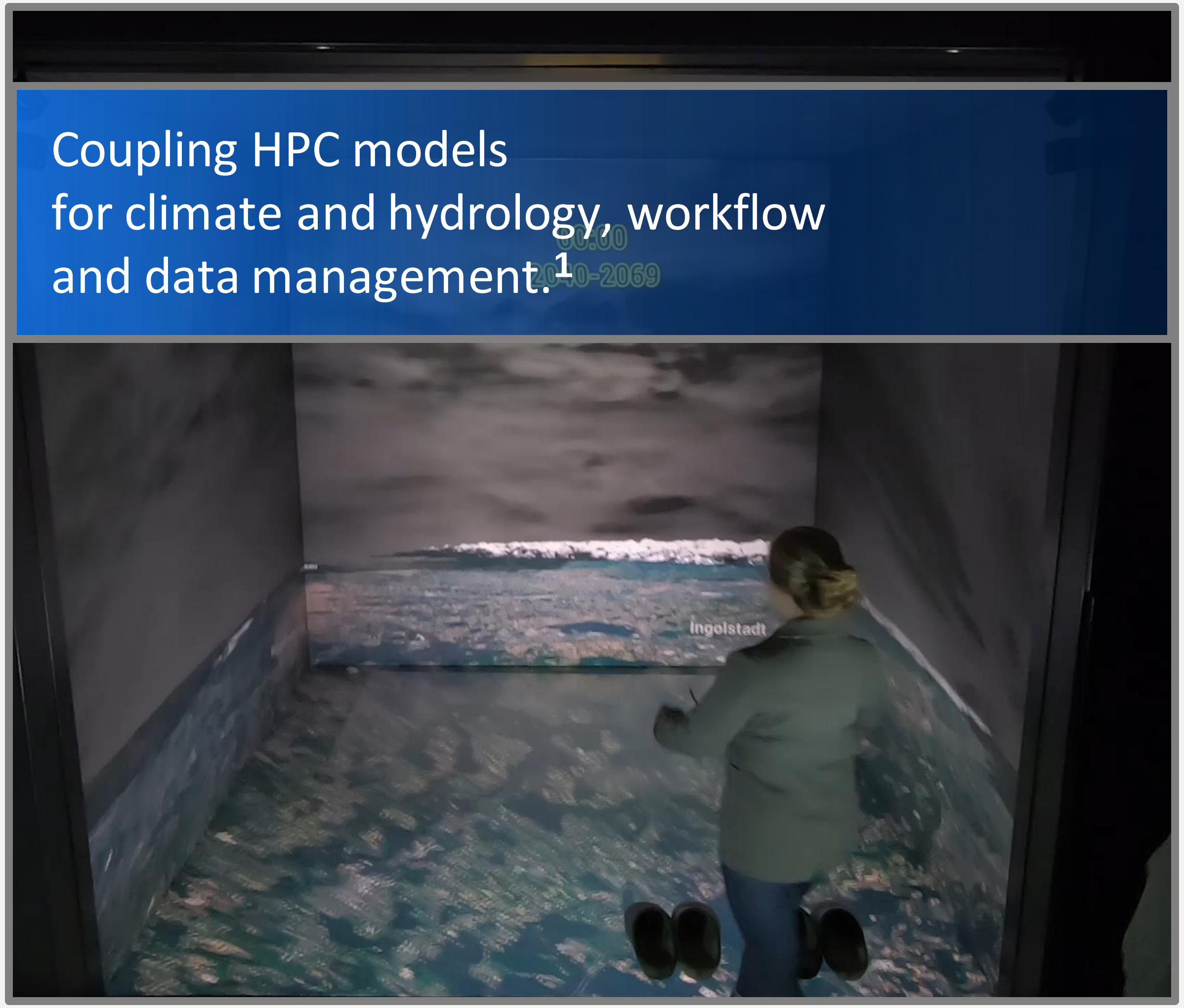


EU Center of Excellence CompBioMed

Personalized medicine: A digital twin of a patient



Bavarian-Canadian Project ClimEx

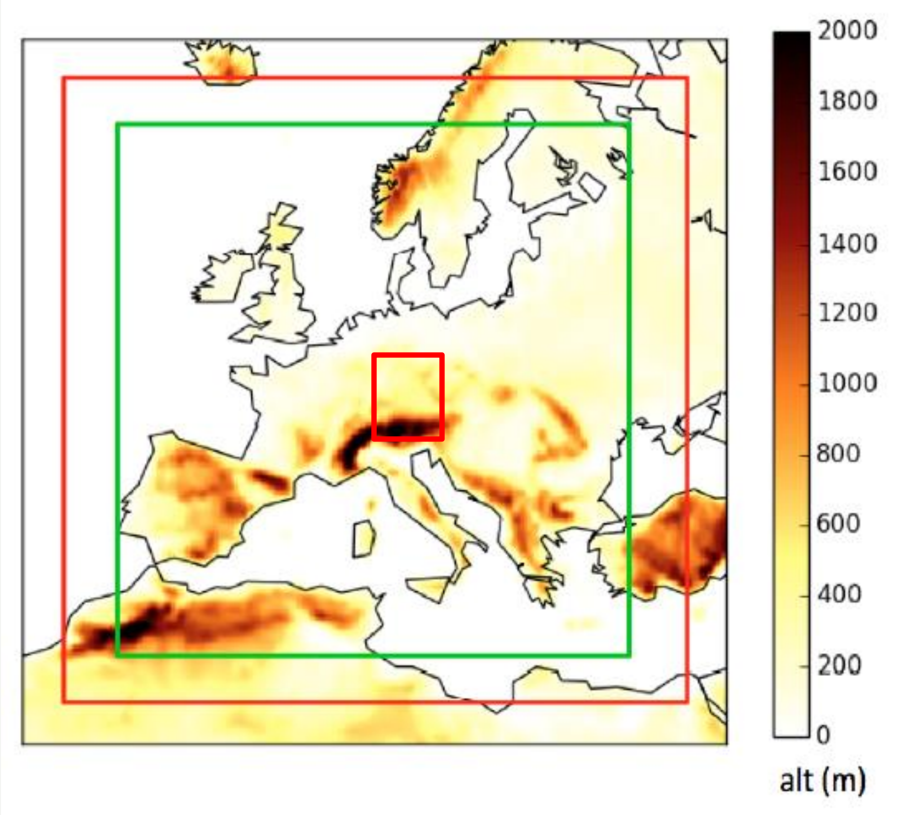


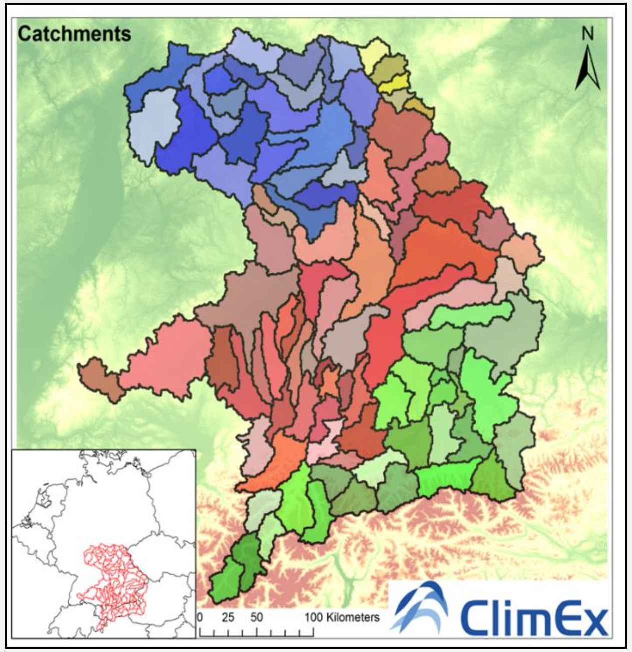
Coupling HPC models for climate and hydrology, workflow and data management.[1]

Impact of the climate crisis on extreme hydro-logical events such as floods and droughts.[2]

www.climex-project.org; **1** © LRZ / Ouranos; **2** Danube Flooding 2013 © sueddeutsche.de

ClimEx – Scales of application



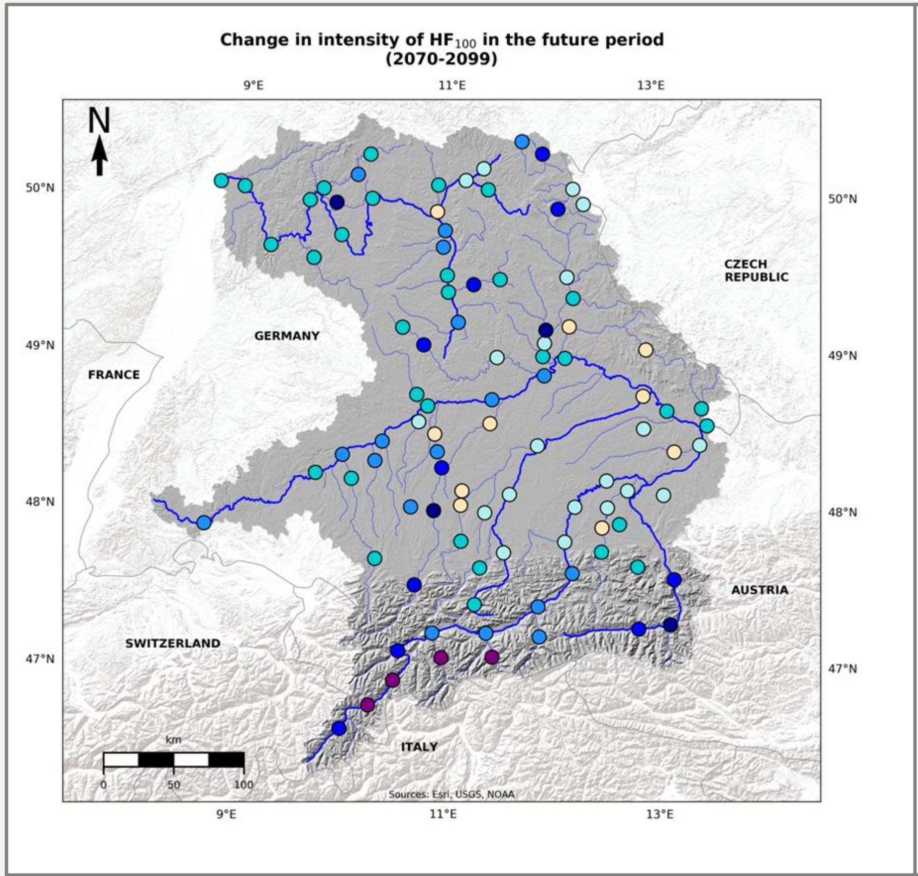


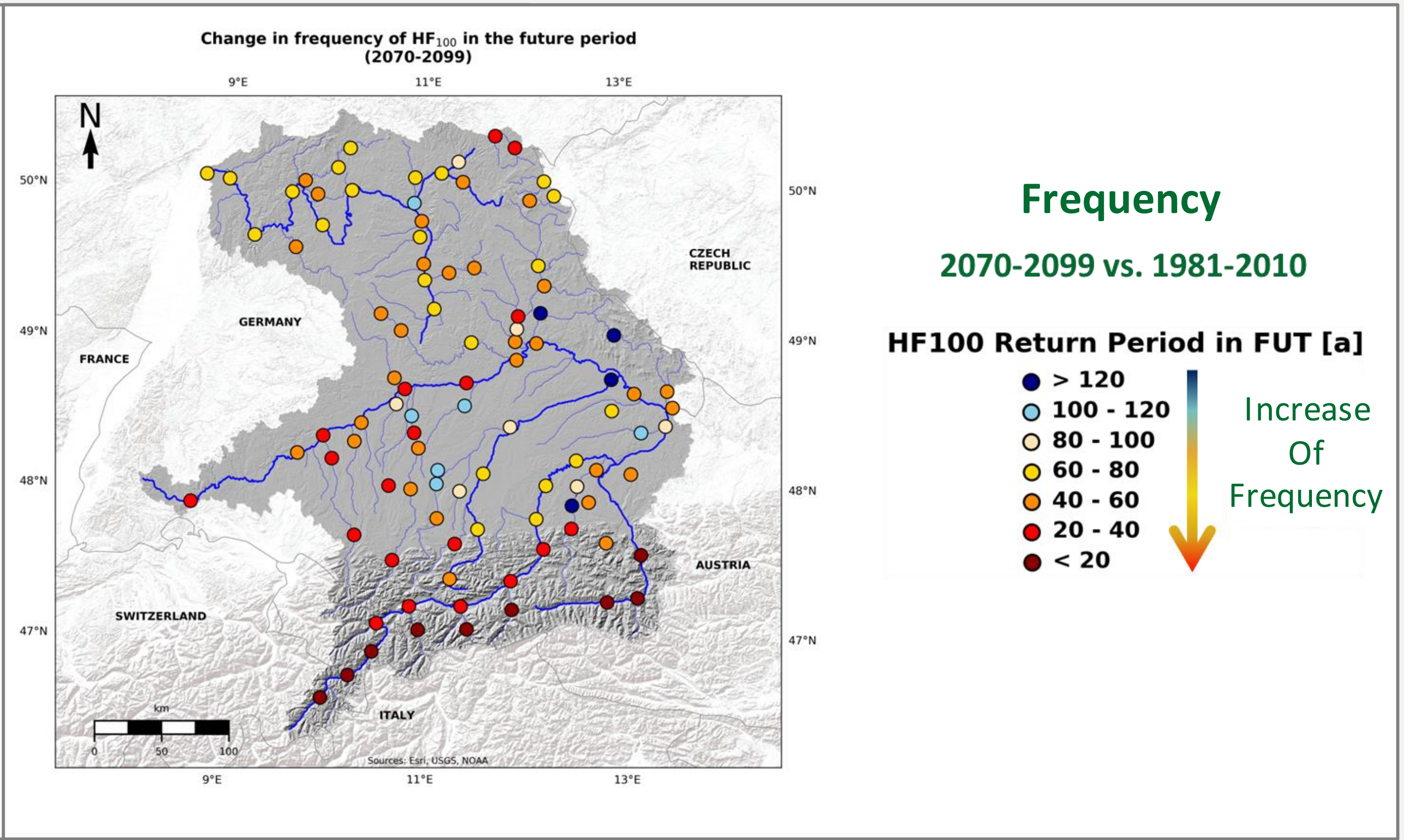
Scope: Assess the CC related dynamics of extremes events from the continental to the river basin scale

Goal: a) Improve process understanding of non-linear hydro-meteorological extreme events
b) Provide adaptation and management options for stakeholders to reduce related risks

Will extreme flows/floods (e.g. HQ100) be more severe?

Changes in Frequency and Intensity of HQ100 in Bavarian river basins





(Willkofer, Ludwig et al., in prep.)

Background...

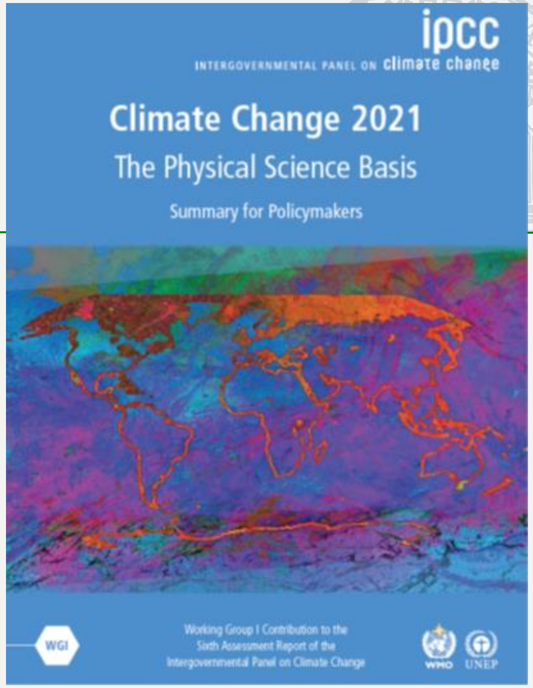
From the 6th Assessment Report of the IPCC (AR6) (2021/2022):
Working Group I: The Physical Science Basis
Working Group II: Impacts, Adaptation and Vulnerability

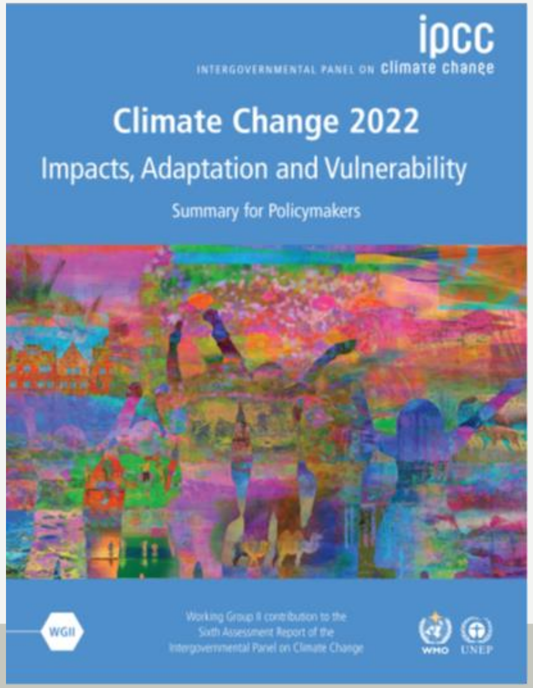
WGI.B – Possible Climate Futures

B.2 Many changes in the climate system become larger in direct relation to increasing global warming. They include increases in the frequency and intensity of hot extremes, marine heatwaves, and heavy precipitation, agricultural and ecological droughts in some regions […]

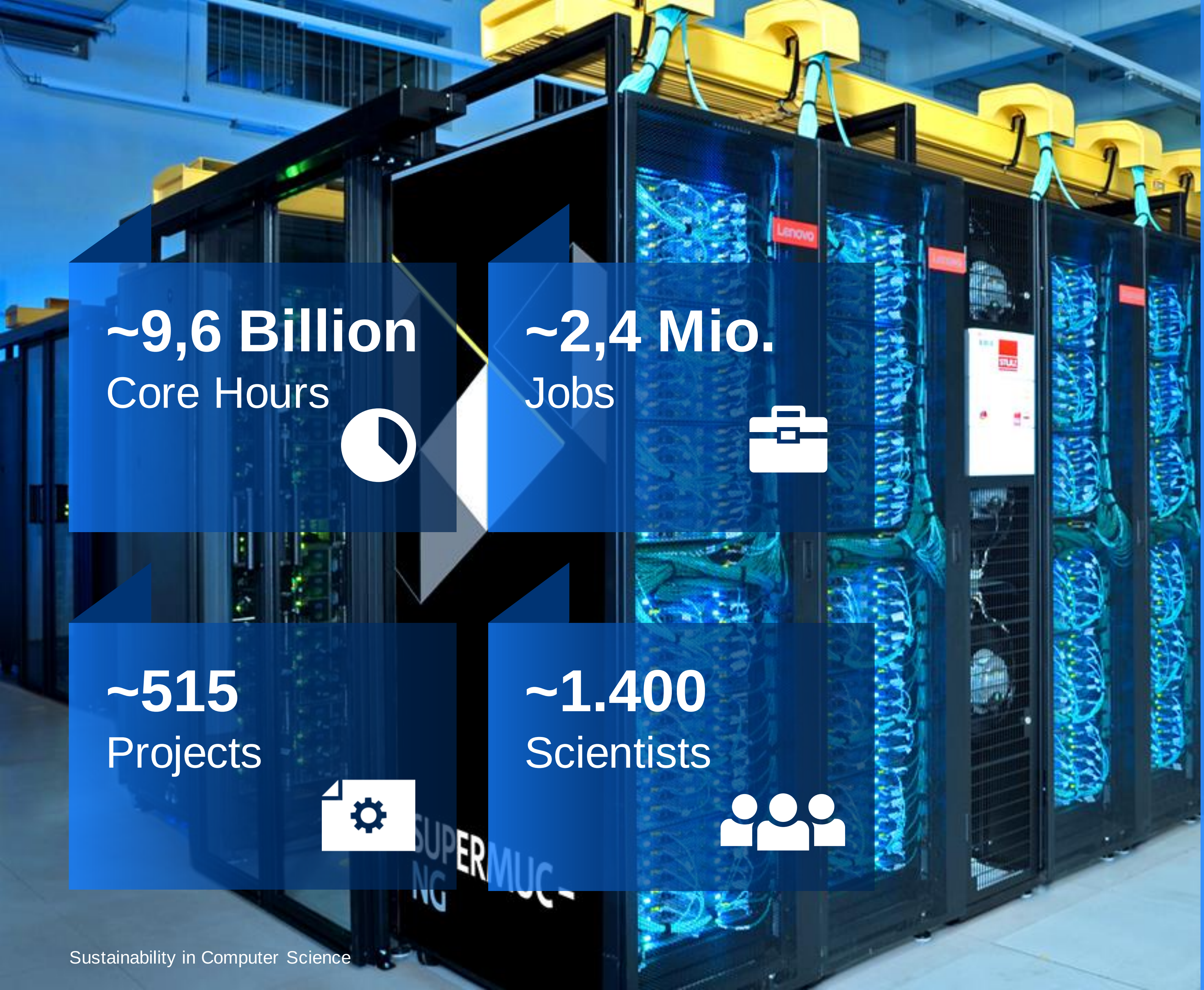
WGII.C – Future Adaptation Options and their Feasibility

C.2 […] The effectiveness of adaptation to reduce climate risk is documented for specific contexts, sectors and regions (high confidence) and will decrease with increasing warming. Integrated, multi-sectoral solutions […] increase the feasibility and effectiveness of adaptation in multiple sectors.





Quelle: F. Löchner für LRZ



~9,6 Billion
Core Hours

~2,4 Mio.
Jobs

~515
Projects

~1.400
Scientists

SuperMUC-NG

Lenovo Intel (2019)

311,040 cores
Intel Xeon Skylake

26.9 PetaFlops Peak

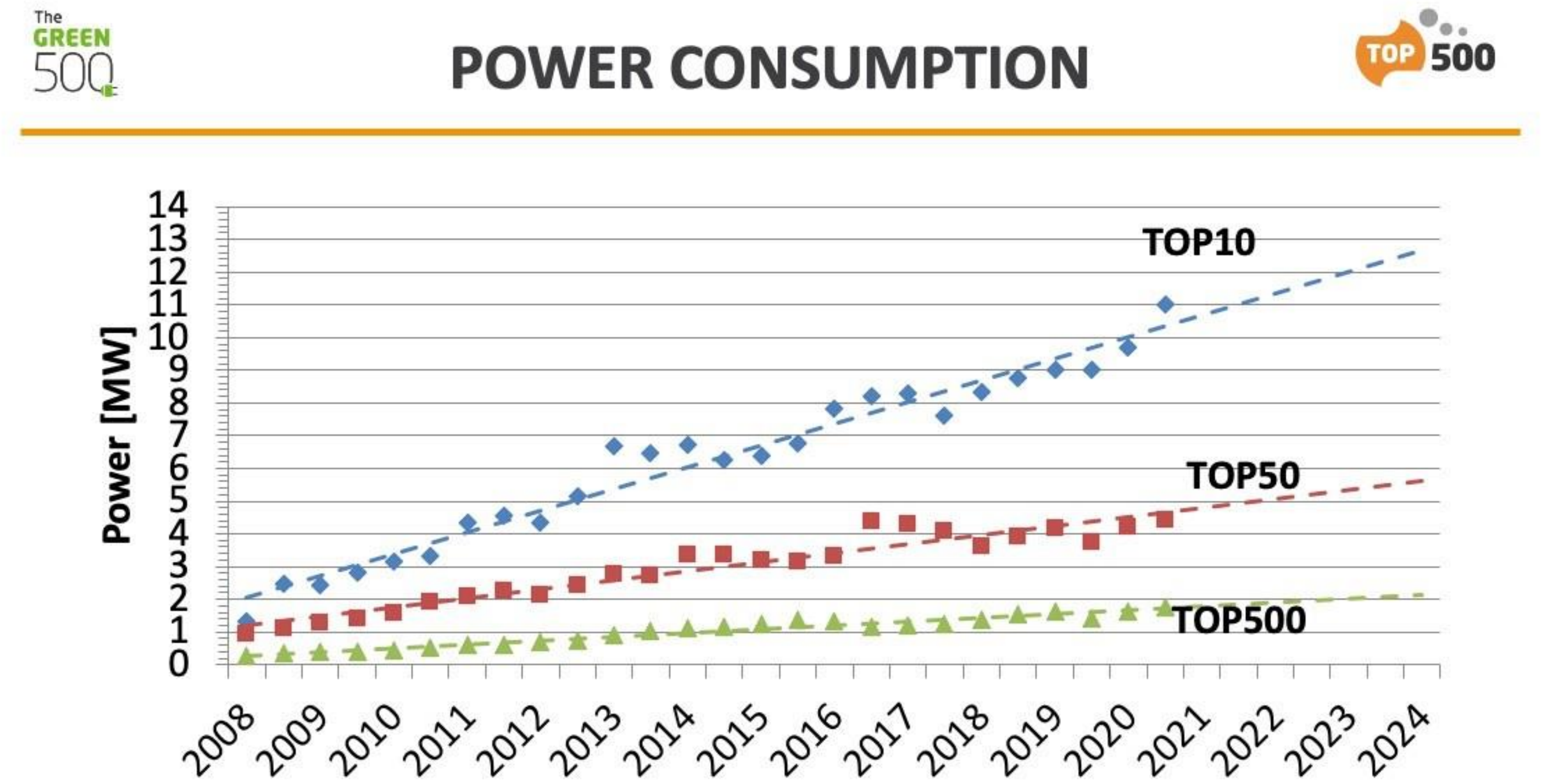
19.5 PetaFlops Linpack*

719 TeraByte Main Memory

70 PetaByte Disk

#
Site
Manufacturer
Computer
Country
Cores
Rmax
Power [MW]
1
Oak Ridge National Laboratory
IBM
Summit
IBM Power System,
P9 22C 3.07GHz, Mellanox EDR, NVIDIA GV100
USA
2,397,824
143.5
9.8
2
Lawrence Livermore National Laboratory
IBM
Sierra
IBM Power System,
P9 22C 3.1GHz, Mellanox EDR, NVIDIA GV100
USA
1,572,480
94.6
7.4
3
National Supercomputing Center in Wuxi
NRCPC
Sunway TaihuLight
NRCPC Sunway SW26010,
260C 1.45GHz
China
10,649,600
93.0
15.4
4
National University of Defense Technology
NUDT
Tianhe-2A
ANUDT TH-IVB-FEP,
Xeon 12C 2.2GHz, Matrix-2000
China
4,981,760
61.4
18.5
5
Swiss National Supercomputing Centre (CSCS)
Cray
Piz Daint
Cray XC50,
Xeon E5 12C 2.6GHz, Aries, NVIDIA Tesla P100
Switzerland
387,872
21.23
2.38
6
Los Alamos NL / Sandia NL
Cray
Trinity
Cray XC40,
Intel Xeon Phi 7250 68C 1.4GHz, Aries
USA
979,072
20.16
7.58
National Institute of Advanced Industrial Science and Technology
Fujitsu
AI Bridging Cloud Infrastructure (ABCI)
PRIMERGY CX2550 M4,
Xeon Gold 20C 2.4GHz, IB-EDR, NVIDIA V100
Japan
391,680
19.9
1.65
Leibniz Rechenzentrum
Lenovo
SuperMUC-NG
ThinkSystem SD530,
Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path
Germany
305,856
19.5
Oak Ridge National Laboratory
Cray
Titan
USA
560,640
17.6
Lawrence Livermore National Laboratory
1,572,864
17.2
7.89
SC18

Average Power Consumption Top 10/50/500 Systems



Source: https://twitter.com/top500supercomp/status/1329583763844599808

LRZ HPC Systems and Energy Efficiency

- 2012: SuperMUC – first hot water cooled system (CPU+Memory)
- 2017: CoolMUC 3 - first 100% hot water cooled system (fan-less)
- 2018: SuperMUC-NG - higher temperatures (45 degree Celsius), 100% free cooling throughout the year, PUE = 1,03
- 2024/25: ExaMUC - ??? (procurement ongoing innovation partnership)

The Four Pillar Model

Neighboring Buildings

Data Center

Improve PUE (Power Usage Effectiveness)	Reduce Hardware Power Consumption	Optimize Resource Usage, Tune System	Optimize Application Performance
Building Infrastructures	**HPC System Hardware**	**HPC System Software**	**HPC Applications**

Holistic Optimization Strategy

Torsten Wilde, Axel Auweter, Hayk Shoukourian
„The 4 Pillar Framework for energy efficient HPC data centers“,
Comput Sci Res Dev (2014) 29: 241. https://doi.org/10.1007/s00450-013-0244-6

LRZ HPC Systems and Energy Efficiency

- 2012: SuperMUC – first hot water cooled system (CPU+Speicher)
- 2017: CoolMUC 3 - first 100% hot water cooled system (fan-less)
- 2018: SuperMUC-NG - higher temperatures (45 degree Celsius), 100% free cooling throughout the year, PUE = 1,03
- 2024/25: ExaMUC - ??? (procurement ongoing innovation partnership)

- DVFS - Dynamic Voltage Frequency Scaling (on average: 68% of nominal frequency)
- Control of power capping / sleep modes
- 100% renewable energies
- Cooling infrastructure optimizations - Glycol free cooling
- Waste heat usage
- Adsorption cooling machines
- Future (unconventional) computing technologies – Accelerators, Quantum, …

Energy Efficiency at LRZ

- Pioneer in direct-liquid hot-water cooling
- Heat reuse (heating, adsorption chillers)
- Energy-aware scheduling
- Broad instrumentation of all components
- DCDB: High frequency, high resolution monitoring and analysis

What about ChatGPT: https://chat.openai.com/

Question: **Is ChatGPT energy efficient?**

Answer:

- *ChatGPT's energy efficiency depends on factors like **hardware**, **model size**, and **deployment**.*
- *Large models can be **energy-intensive**.*
- *Researchers are working on techniques to improve efficiency.*
- *The actual efficiency varies based on **implementation** and **infrastructure**.*
- *Improvements are expected over time.*

Question: **What is the energy consumption for training ChatGPT?**

Answer:

- *To provide a rough estimate, training GPT-3, which has 175 billion parameters, could consume hundreds of megawatt-hours (MWh) of electricity during the training process.*

Some numbers from research online - Estimates

ChatGPT-3 Energy consumption:

- to train the GPT-3 model: 1064 MWh
- To operate search requests: 260 MWh/day
- A ChatGPT query: 6,79 Wh
- https://medium.com/@zodhyatech/how-much-energy-does-chatgpt-consume-4cba1a7aef85#

Comparison:

- With 1 MWh, an electric car can travel about 5000 km.
- The average per capita electricity consumption in Germany is about 1.5 MWh/a,
  per household about 4 MWh/a.
- 260 MWh/d correspond to the power consumption of a small town with 24000 households or 64000 inhabitants.

www.top500.org

Top 500 Supercomputers

- List: June 2023
- Vendor: NVIDIA
- Top 5 out of 16 entries

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
9	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63.46	79.22	2,646
14	**Pre-Eos 128 Node DGX SuperPOD** - NVIDIA DGX H100, Xeon Platinum 8480C 56C 2GHz, NVIDIA H100 Tensor Core GPUs, NVIDIA ConnectX-7 NDR 400G Infiniband, Nvidia NVIDIA Corporation United States	81,920	40.66	58.05	
40	**HiPerGator AI** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Infiniband HDR, Nvidia University of Florida United States	138,880	17.20	21.31	583
55	**Christofari Neo** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 80GB, Infiniband, Nvidia SberCloud Russia	98,208	11.95	14.91	
58	**KT DGX SuperPOD** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Infiniband HDR, Nvidia Korea Telecom South Korea	98,208	10.38	14.42	

Top 500 Supercomputers for HPC and Artificial Intelligence

NVIDIA Technology:

- DGX A100
  https://www.nvidia.com/en-us/data-center/dgx-a100/
- NVIDIA DGX H100
  https://www.nvidia.com/en-us/data-center/dgx-h100/
- NVIDIA DGX SuperPod
  https://www.nvidia.com/en-us/data-center/dgx-superpod/

- All current systems are air cooled
- Water cooling is only a supplement
- PUE: 1,65-1,80
  - +30% Fans
  - +30% Cooling Infrastructure

The Munich Center for Machine Learning (MCML)

Left:
- 2-3 NVIDIA DGX A100
- =16-24 GPUs per Rack
- Air-cooled
  (in a water-cooled rack)
- PUE: 1,65-1,80

Power Usage Effectiveness:

the **total amount of power** entering a data center divided by the power used to run the **IT equipment**

Example:

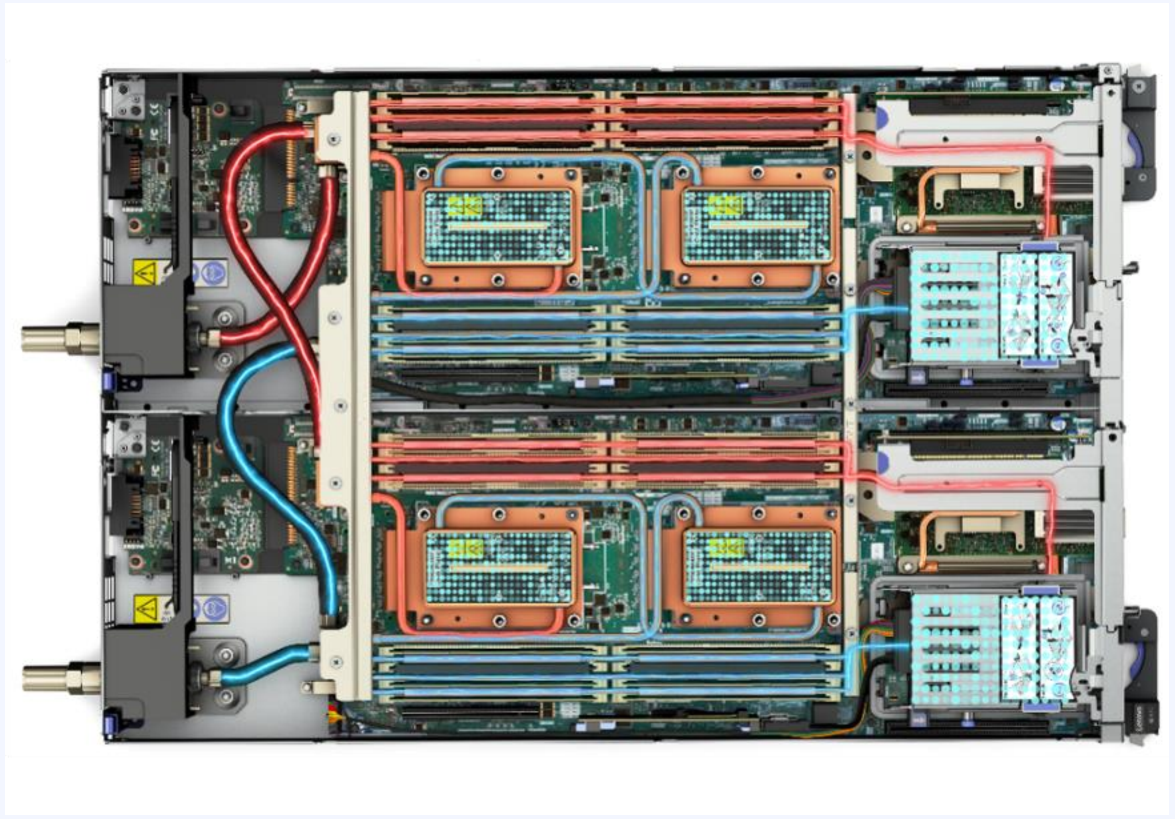
PUE 1,65 = 1 für IT
+ 65% für Infrastructure

Air vs. Water

Name	Unit	Air	Water	Factor
Thermal Conductivity	W/(mK)	0.026	0.598	23x
Heat Capacity	J/(gK)	1.006	4.185	4x
Volumetric Heat Capacity	$kJ/(m^3K)$	1.196	4178	3493x
Thermal Inertia	$J/(m^2Ks^{1/2})$	5.563	1581	284x

Hot water cooling in the compute room

SuperMUC-NG Node – with hot water cooling



Source: Photos Lenovo – Produkt: https://www.lenovo.com/de/de/data-center/servers/high-density/ThinkSystem-SD650/p/77XX7DSSD65

SuperMUC-NG Cool Manager

3.000 l/h pro Rack

Reduced server power consumption

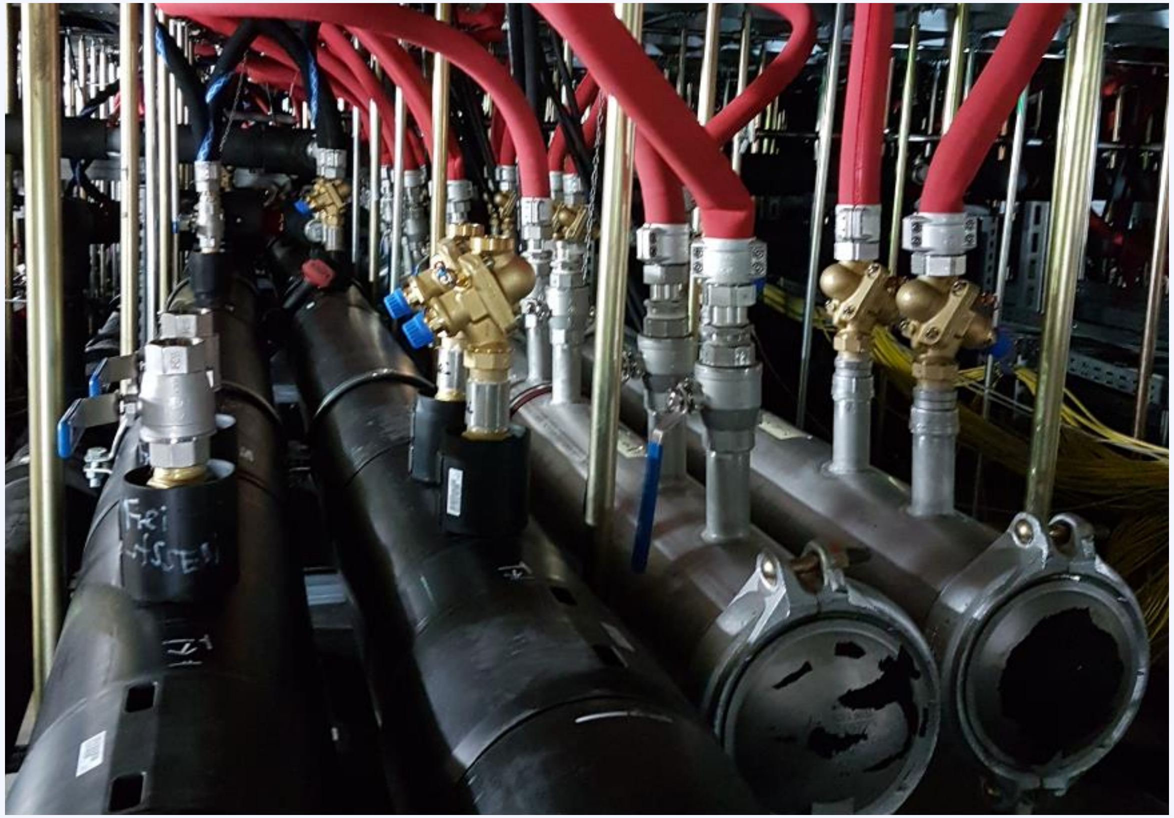
- Lower processor power consumption (~ 5%)
- No fan per node (~ 4%)

Reduce cooling power consumption

- At 40°C free cooling all year long (> 15%)

Energy Aware Scheduling

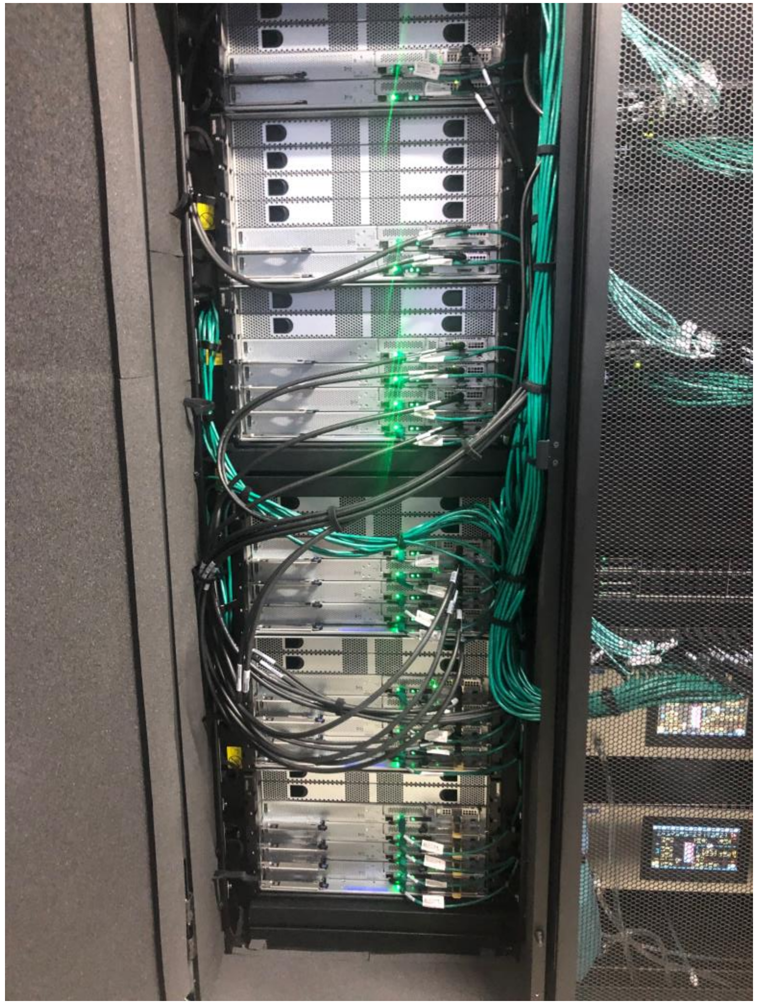
- Only CPU bound jobs get max frequency (~ 5%)



The Munich Center for Machine Learning (MCML)

Left:
- 2-3 NVIDIA DGX A100
- =16-24 GPUs per Rack
- Air-cooled
  (in a water-cooled rack)
- PUE: 1,65-1,80

Right:
- Up to 144 GPUs per Rack
- Lenovo HGX-based nodes (4x A100)
- Direct hot-water cooled (allows free cooling)
- PUE: 1,03-1,05



SuperMUC-NG Phase 2

Lenovo Intel

240
Direct hot water cooled compute nodes (CPU:Sapphire Rapids + GPU:Ponte Vecchio)

SD650-I v3
Lenovo platform

1 PB
DAOS storage system mit Intel Xeon Scalable CPUs and 3rd Gen Intel® Optane™ persistent memory


HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg
hpc.bw
HSUper
576 MEGWARE Compute Nodes
5 MEGWARE GPU Nodes
Direct-to-chip liquid cooling
1162 Intel® Xeon Scalable 8360Y
41,832 CPU cores
149 TB DDR4 main memory
3.21 PFlop/s RPEAK (CPU-only)
NVIDIA® HDR InfiniBand Network
1 PB net BeeGFS Storage System
1 PB net Ceph Storage System
MEGWARE

Savings potentials through direct water cooling

	Infrastructure		IT	
Investment	↓	Fewer chillers requires	↗	IT slightly more expensive (depending on quantity)
Operations costs	↓	Power consumption for refrigeration is significantly reduced	↓	Power consumption of water-cooled IT is lower (fewer fans)

Using the waste heat – heating and adsorption cooling

Foto: Connie Zhou/Google/Zuma https://www.nature.com/articles/d41586-018-06610-y

Example

Google‘s Water Demand in Luxembourg

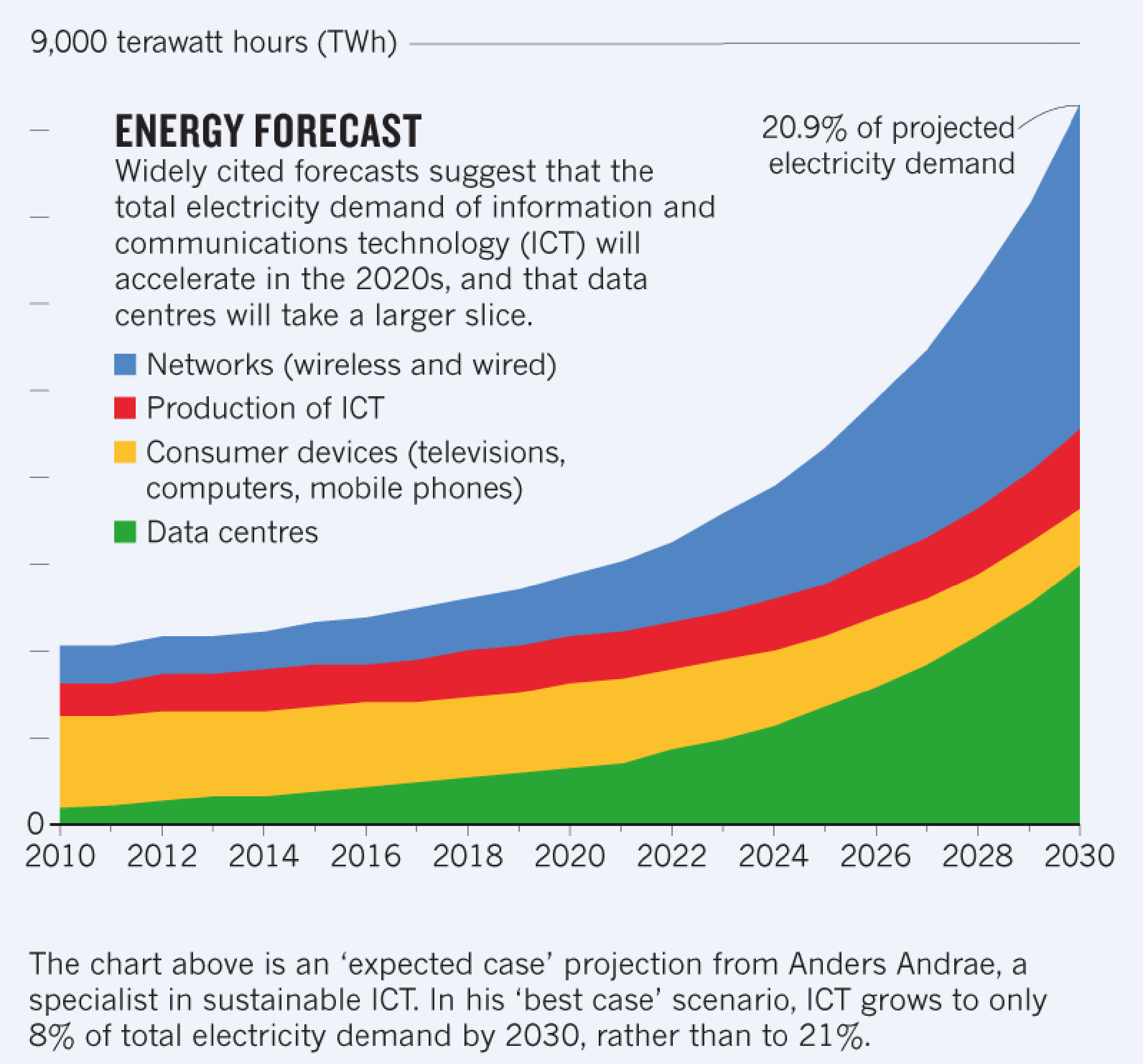
- New Google data center in Luxembourg - €1.2 billion
- 12% of Luxembourg‘s electricity demand
- Cooling water demand: 10 million l/day
- Google refuses to provide information on the grounds that it "*could reveal details of the DC and the agreement with the Luxembourg government to competitors*"

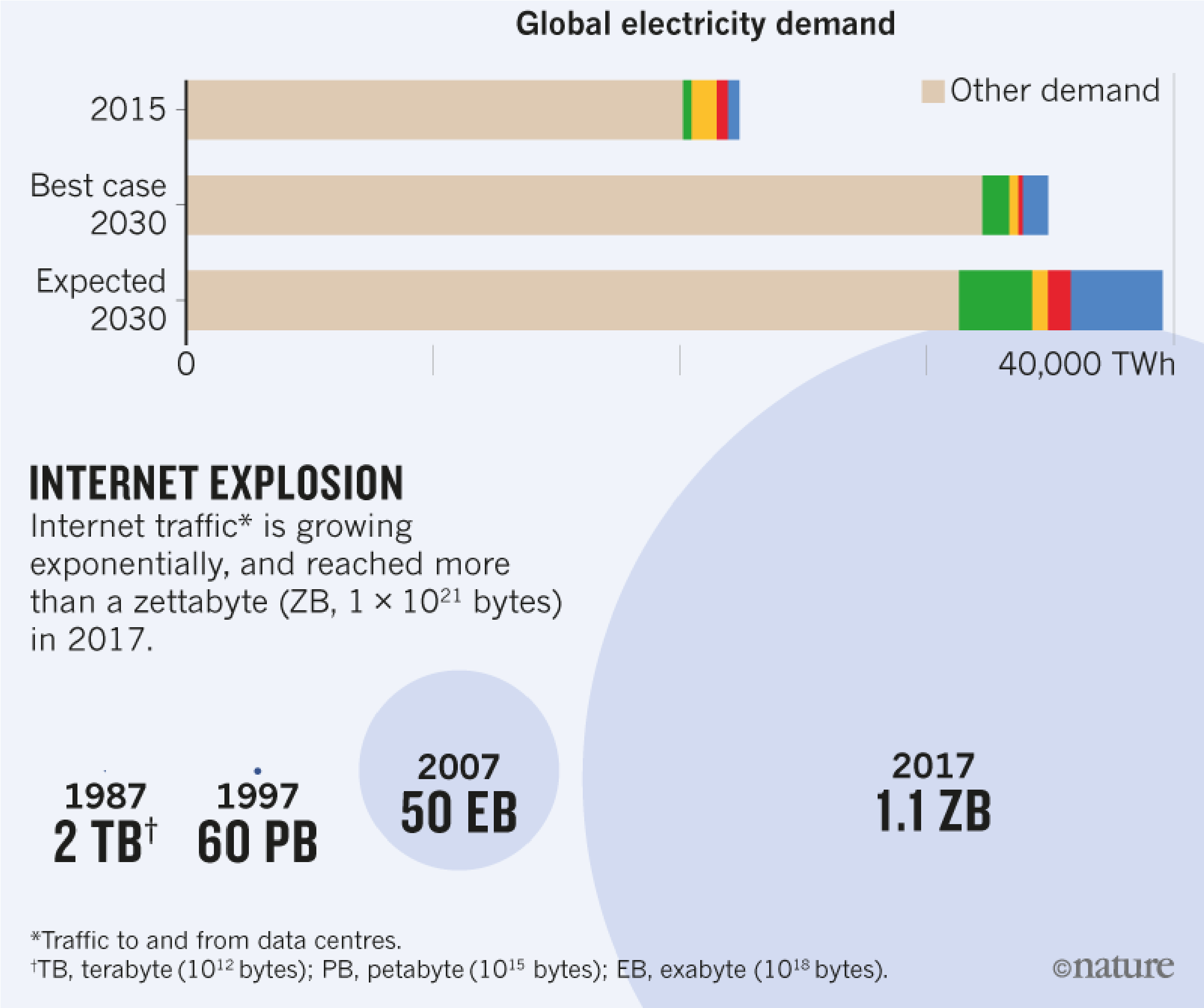
https://orf.at/stories/3184229/

Foto: Reuters/Yves Herman

Nature News Feature - https://www.nature.com/articles/d41586-018-06610-y

How to stop data centres from gobbling up the world’s electricity

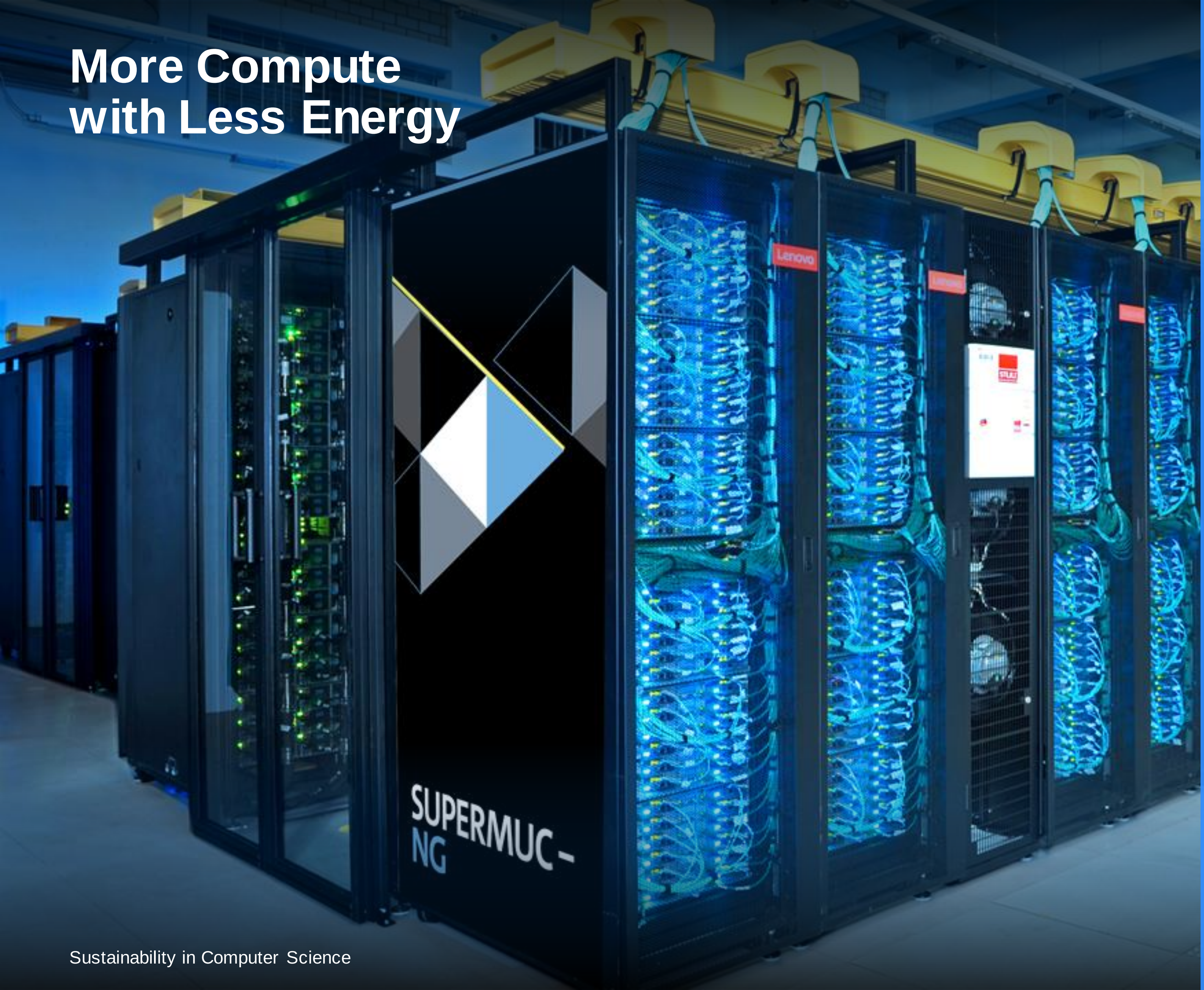




Proposed German government bill EnEfG 2023

- Section 5 - Energy efficiency in data centers
- § 23 Energy efficiency and waste heat requirements for data centers
- (1) Data centers that begin operations on or after January 1, 2025, must, for the first two years from the start of operations
  - a planned power usage effectiveness (PUE) **smaller or equal to 1,3**
  - a planned percentage of **reused energy of at least 30 percent**; by January 1, 2027, of at least 40 percent.
- …
- (3) For data centers that begin operations on or after January 1, 2024, the **minimum inlet** temperature for air cooling of information technology **is 27 degrees Celsius**; a lower inlet temperature is only permitted if it can be achieved without the use of a refrigeration system.

More Compute with Less Energy



- **Energy efficiency** is essential for sustainability
- We need a **holistic approach** - more than GFlop/watt
- Aside from the **carbon footprint**, there is an important **financial aspect**

Contact: Prof. Dr. Dieter Kranzlmüller
kranzlmueller@lrz.de | @Kranzlmueller