Less Moore, More Brain

Меньше Мура, Больше Мозга

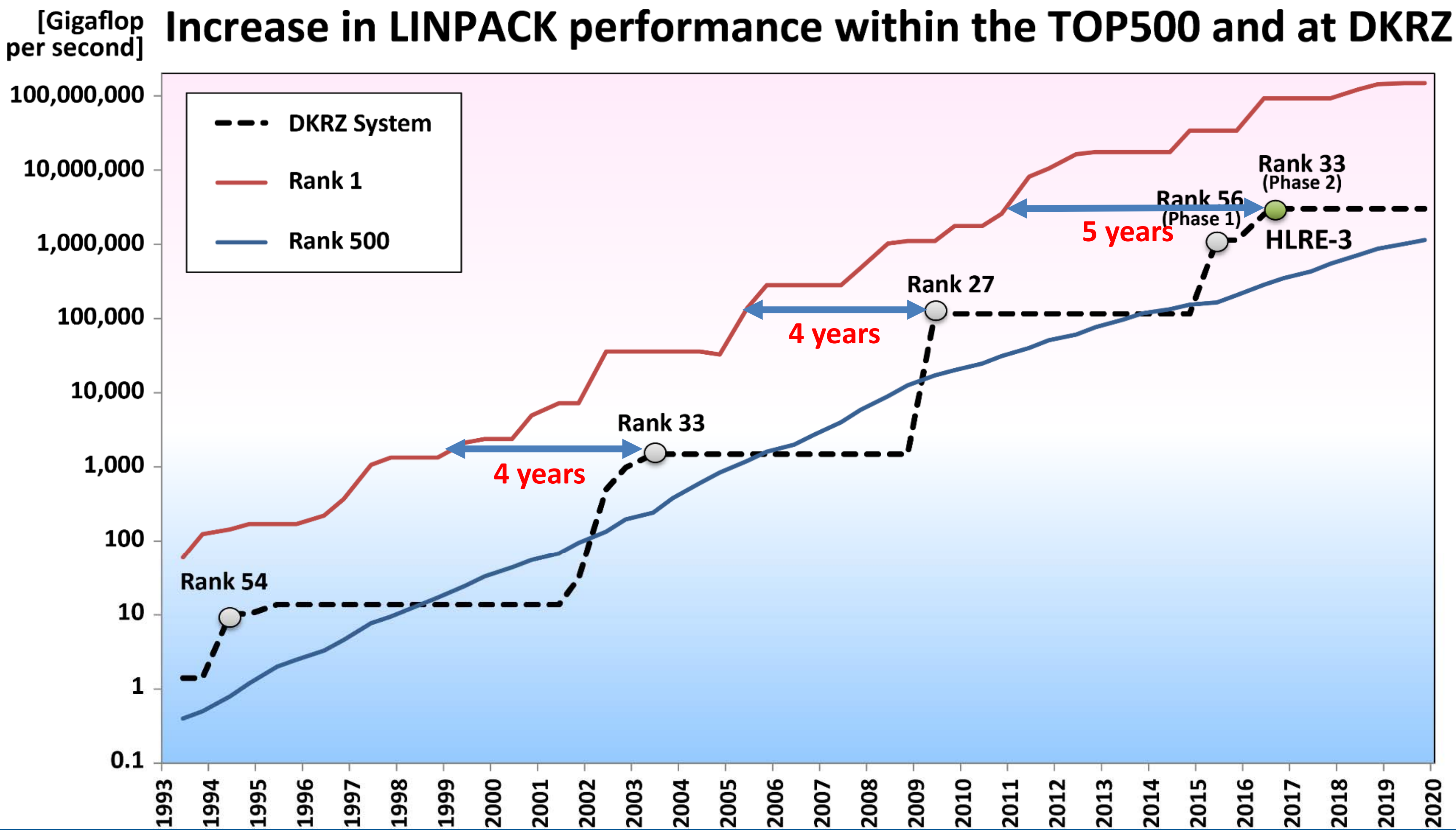
Prof. Dr. Thomas Ludwig
German Climate Computing Center (DKRZ)
University of Hamburg, Department for Computer Science (UHH/FBI)

Less Moore

Меньше Мура

Powerful Computers for Competitive Research

Increase in LINPACK performance within the TOP500 and at DKRZ



HLRE-3 – Mistral (2015-2020)

bullx DLC 720

3,500+ nodes, 100,000+ cores, Haswell/Broadwell, 3.6 PFLOPS

54 PB disk storage with Lustre

no GPUs for acceleration or ML

TOP500 List in November 2019

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
2	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
3	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
4	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482

factor 7x to ExaFLOPS
Summit 4 lists on #1

TOP500 List in June 2020

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,299,072 **3.0x**	415,530.0 **2.8x**	513,854.7 **2.5x**	28,335 **2.8x**
2	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371

TOP500 Positionings of Blizzard and Mistral

List Entries	Rank Blizzard	Rank Mistral
06/2009 06/2015	27	(56)
11/2009 11/2015	35	(64)
06/2010 06/2016	41	34
11/2010 11/2016	59	35
06/2011 06/2017	72	39
11/2011 11/2017	100	43
06/2012 06/2018	154	55
11/2012 11/2018	234	62
06/2013 06/2019	371	73
11/2019	out	80
06/2020	out	93

extrapolation of Bull during ISC 2016

„this machine will not fall out of the list during the 5 years of operation!“

Extrapolation since 1993 - **worst case** scenario

2 600 TFLOPS	date	Rank
First presence in TOP500 List	Jun-16	32
Last presence in TOP500 List	Nov-20	492

Extrapolation from last 5 years - best case scenario

2 600 TFLOPS	date	Rank
First presence in TOP500 List	Jun-16	32
Last presence in TOP500 List	Nov-20	417

Moore´s Law

Moore's law is the observation that the number of transistors in a dense integrated circuit doubles about every two years.

The observation is named after Gordon Moore, the co-founder of Fairchild Semiconductor and CEO of Intel, whose 1965 paper described a doubling every year in the number of components per integrated circuit, and projected this rate of growth would continue for at least another decade.

https://en.wikipedia.org/wiki/Moore%27s_law

Later there are slightly revised versions of the law.

Dennard Scaling

Dennard observes that transistor dimensions could be scaled by 30% (0.7x) every technology generation, thus reducing their area by 50%. This would reduce circuit delays by 30% (0.7x) and therefore increase operating frequency by about 40% (1.4x). Finally, to keep the electric field constant, voltage is reduced by 30%, reducing energy by 65% and power (at 1.4x frequency) by 50%. Therefore, in every technology generation, if the transistor density doubles, the circuit becomes 40% faster, and power consumption (with twice the number of transistors) stays the same.

https://en.wikipedia.org/wiki/Dennard_scaling

End of Dennard Scaling

Historically, the transistor power reduction afforded by Dennard scaling allowed manufacturers to drastically raise clock frequencies from one generation to the next without significantly increasing overall circuit power consumption.

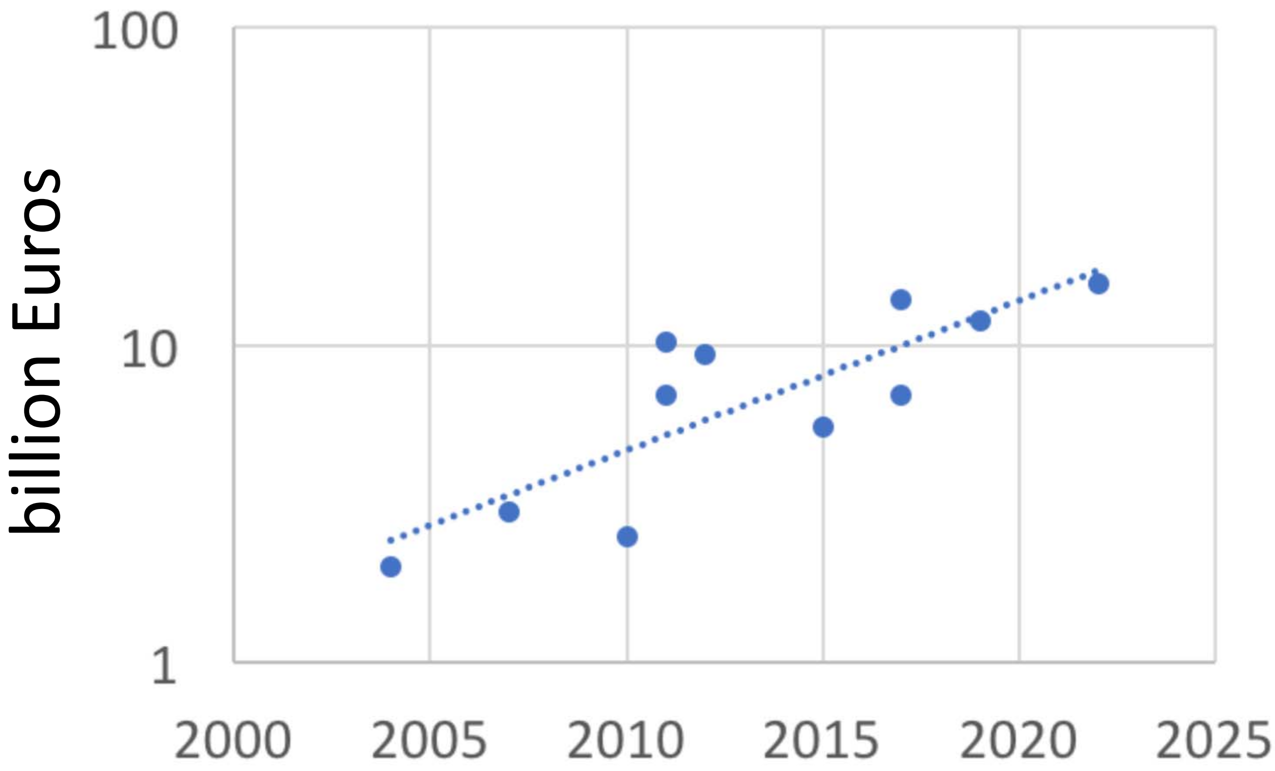
Since around 2005–2007 Dennard scaling appears to have broken down. As of 2016, transistor counts in integrated circuits are still growing, but the resulting improvements in performance are more gradual than the speed-ups resulting from significant frequency increases. The primary reason cited for the breakdown is that at small sizes, current leakage poses greater challenges and also causes the chip to heat up, which creates a threat of thermal runaway and therefore further increases energy costs.

https://en.wikipedia.org/wiki/Dennard_scaling

In Summary

- Frequencies cannot go higher because of too much power consumption and heat emission
- Currently, you still can put more transistors on a chip
- With more transistors you can make more processor cores
- You must decrease the size of the transistors, but…

Chip Factory Costs



Intel, Samsung and TSMC press release [The Economist 2016], [Gartner 2016]

The Deceleration of Acceleration

- Systems at DKRZ
  (same investment in € plus inflation)
  - 2009: NEC -> IBM with 100x improvement
  - 2015: IBM -> Bull/Atos with 22x improvement
  - 2021: Bull/Atos -> Bull/Atos with 5x improvement
    - 2x power consumption
- Partner centers recently
  - LRZ (Garching/München): 4x
  - ECMWF (Bologna): 5x

From our Request for Money in 2016

- DKRZ collected 11 use cases of its users
- Requirements for increase in computational performance were between 7 and 1,000
- Our estimate in 2016 was to achieve a 10x increase
- Instead we will receive a 5x increase

- Consequence: not a single use case can be satisfied!

The End of “Traditional” HPC

Traditional HPC

- Mathematics is differential equations
- Processors are conventional homogeneous CPUs
- Moore´s Law comes to an end
  - economically reasonable chip improvements very difficult

- No more low hanging fruits!
- No more fruits?

Instead: Invest in Brainware

- Better codes, better software engineering
- Adaptations to heterogeneous hardware
- Interdisciplinary teams of domain scientists, computer scientists, mathematicians, …
- Adopt new methods
  - Data intensive science
  - Machine learning (artificial intelligence)

More Brain

Больше Мозга

AI/ML Exceeds Moore´s Law

Anthony Sarkis: "Why AI progress is faster than Moore's Law — the age of the algorithm" (Sep 2018)

https://medium.com/@anthony_sarkis/the-age-of-the-algorithm-why-ai-progress-is-faster-than-moores-law-2fb7d5ae7943

- deep learning algorithms improvement
  - region proposal network
  - image classification
- specialized hardware (GPGPU,…)

Region Proposal Network

- a sub system used in many object detection systems
- “mAP” is a measure for network quality

	R-CNN	Fast R-CNN	Faster R-CNN
Test time per image (with proposals)	50 seconds	2 seconds	**0.2 seconds**
(Speedup)	1x	25x	**250x**
mAP (VOC 2007)	66.0	**66.9**	**66.9**

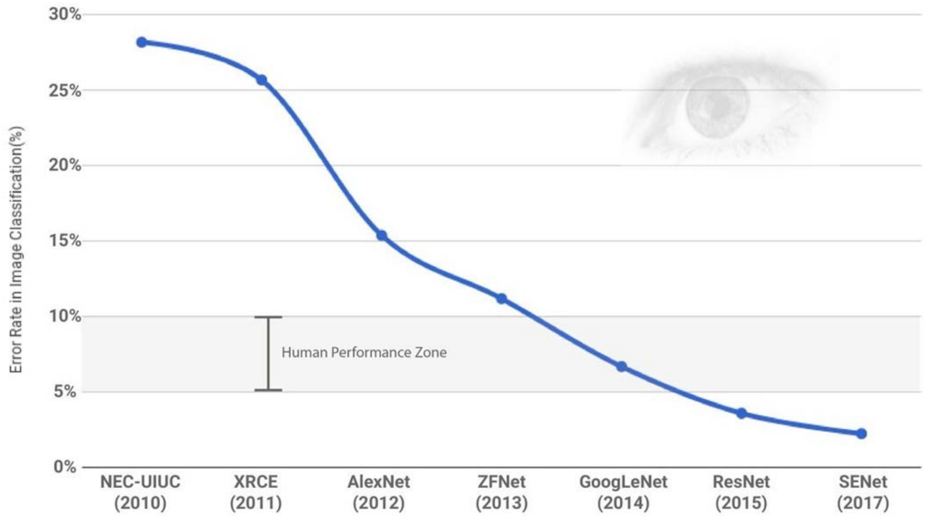
- time to improve from “fast” to “faster”: <1 year

https://arxiv.org/abs/1504.08083

Image Classification

- improvement in image classification
  - Large Scale Visual Recognition Challenge (LSVRC)

    http://www.image-net.org/challenges/LSVRC/



- AlexNet in 2012 uses deep convolutional networks

Climate Science with Machine Learning Gains Momentum!

Exascale Deep Learning for Climate Analytics

Kurth, Treichler, Romero, Mudigonda, Luehr, Phillips, Mahesh,
Matheson, Deslippe, Fatica, Prabhat, Houston

2018

Specific Types of Contributions

ACM Gordon Bell Prize

Innovations in applying high-performance computing to science, engineering, and large-scale data analytics

We extract pixel-level masks of **extreme weather patterns** using variants of Tiramisu and DeepLabv3+ **neural networks**. We describe improvements to the software frameworks, input pipeline, and the network training algorithms necessary to efficiently scale deep learning on the Piz Daint and Summit systems. The Tiramisu network scales to **5300 P100 GPUs** with a sustained throughput of **21.0 PF/s** and parallel efficiency of 79.0%. DeepLabv3+ scales up to **27360 V100 GPUs** with a sustained throughput of **325.8 PF/s** and a parallel **efficiency of 90.7%** in **single precision**. By taking advantage of the **FP16 Tensor Cores**, a **half-precision** version of the DeepLabv3+ network achieves a peak and sustained throughput of 1.13 EF/s and 999.0 PF/s respectively.

NCAR Workshop in Stresa, Italy (Sep. 2019)

https://www2.cisl.ucar.edu/events/conferences/icas/2019

- "Applying Data Analytics and Machine Learning to Storage Systems to Help Meet Organizational Goals“, Bill Anderson, **NCAR**
- "How to use deep learning in weather and climate models“, Peter Dueben, **ECMWF**
- "Machine Learning and Data Driven HPC at NCAR: Strategy and State of Play“, Rich Loft, **NCAR**
- "Insights into Cloud Technologies and Machine Learning for Scientific Data“, Sara Graves, **University of Alabama**, Huntsville
- "Machine Learning at the Met Office“, Niall Robinson, **UK Met Office** Informatics Lab

ECMWF organized workshop on ML and climate/weather

A Comment on Weather, Climate, and ML

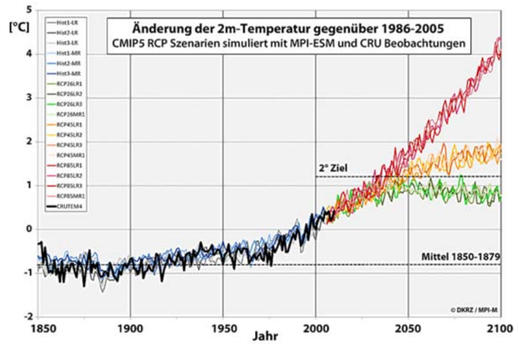
- Weather

  Weather is the **state of the atmosphere**

- Climate

  Climate is the **statistics of weather** over a usually 30 years interval

- Computational weather prediction
  - Make it quick
- Computational climate projection
  - Make it exact – a challenge for ML



AI/ML Research Group at DKRZ

Climate Informatics and Technologies

- Interface between AI/ML and climate science
- Utilization of cutting-egde AI/ML technologies for climate scientists
- Focus on research challenges like e.g. climate prediction and missing climate information of the past

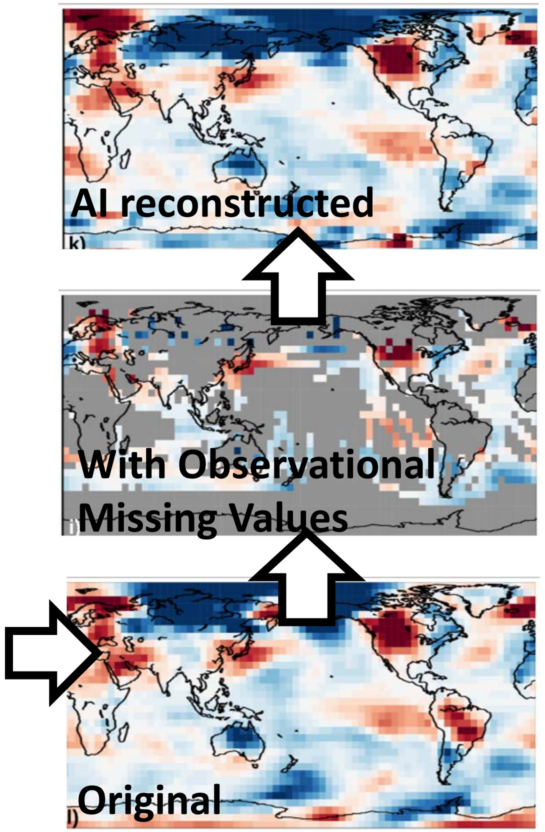
Climate Models

Machine Learning

Example: Missing Historical Climate Information in Observations

Nvidia Technology: Image inpainting on irregular holes using deep convolutional neural networks

Kadow, Hall, Ulbrich: https://www.nature.com/articles/s41561-020-0582-5



Gray -> Missing Values

Retrospect and Prospect

- Machine learning and specialized hardware have a 60+ years success record
- Traditional HPC slows down because of HW issues
- ML is now being adopted by classical HPC communities

- With computational climate science we still explore the application fields
- DKRZ will extend its methods and services portfolio