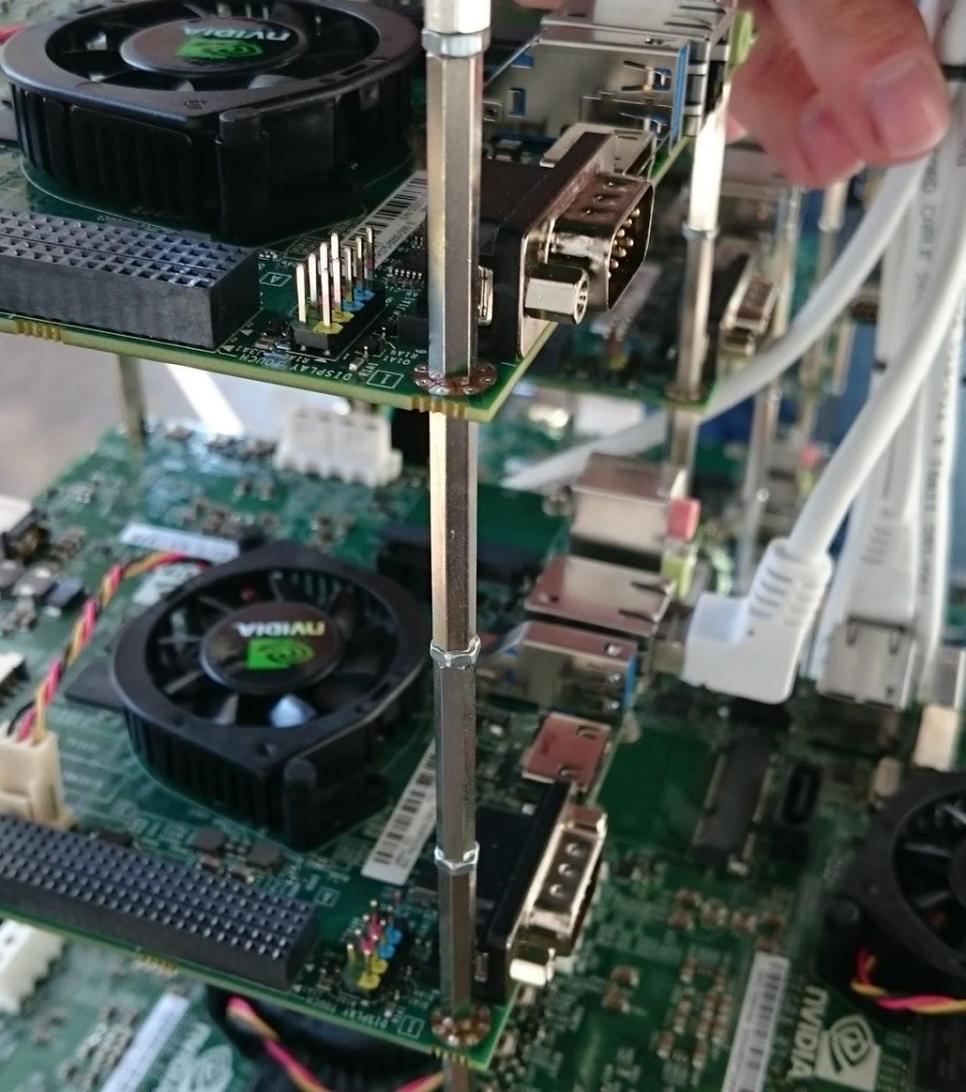




**Mit Bastelrechnern
gegen die
Energiekrise**

Markus Geveler

@ NerdNite 2017 Magdeburg



Markus @ TUDo

www.mathematik.tu-dortmund.de/lisiii/_Markus_Geveler

markus.geveler@math.tu-dortmund.de

Energieeffizientes HPC,
Gebietszerlegungs- und
Mehrgitterverfahren,
Strömungssimulation

tu technische universität
dortmund

fakultät für
mathematik **m!**

nvidia

GPU
EDUCATION
CENTER



I.C.A.R.U.S.

Ministerium für Innovation,
Wissenschaft und Forschung
des Landes Nordrhein-Westfalen



www.icarus-green-hpc.org

Insular
Compute center
Applied Sciences
Renewables
Unconventional
System Integration

@ **TU Dortmund:**
Markus Geveler, Stefan Turek,
Dirk Ribbrock, Hannes Ruelmann,
Daniel Donner, uvm.
@ **MPI Magdeburg**
Peter Benner, Jens Saak,
Martin Köhler, Gerry Truschkewitz

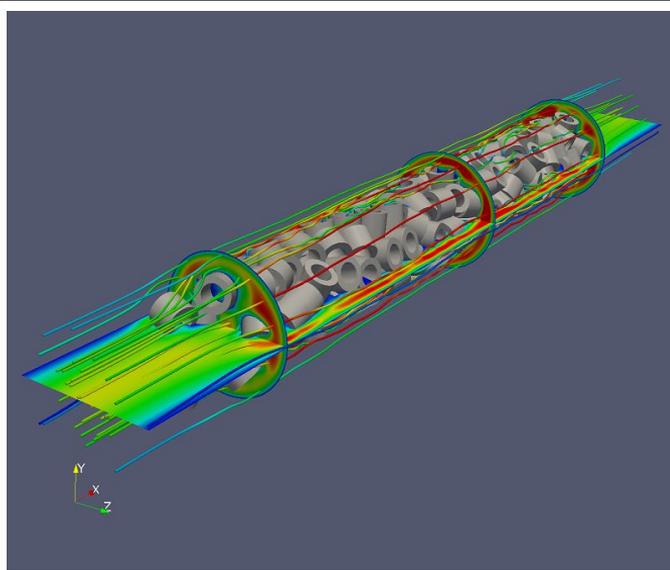


1

(Scientific) Computing in der Energiekrise

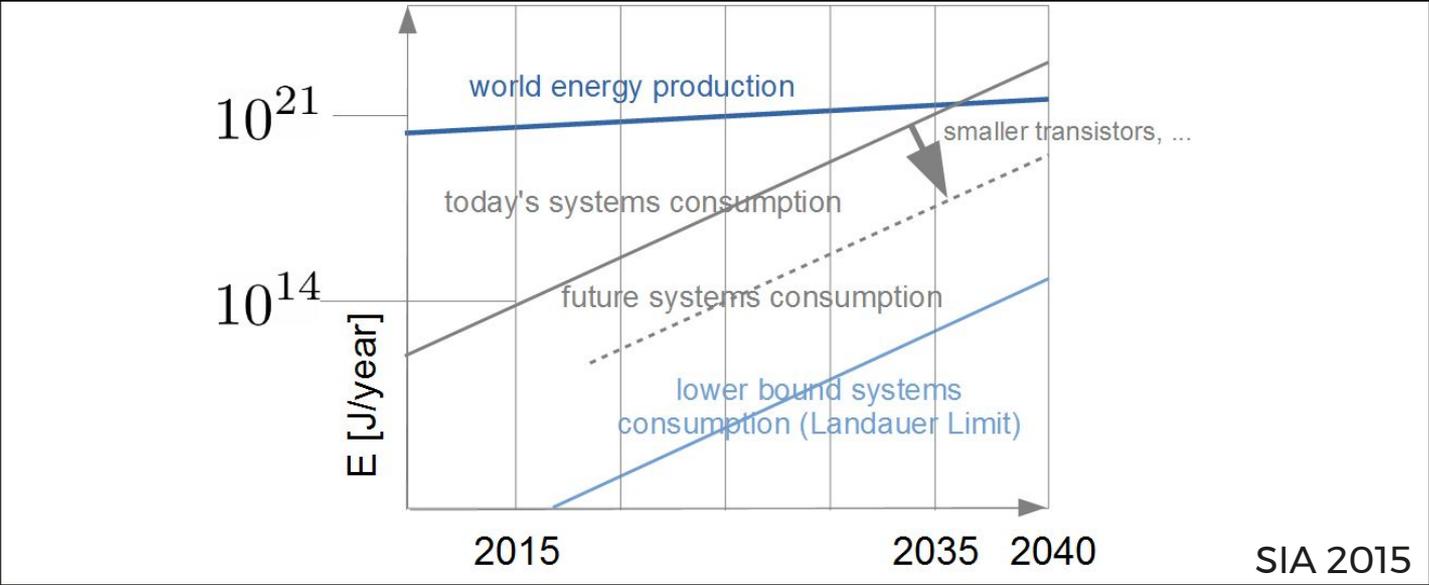
Warum wir nicht so weiter machen können, wie bisher

Wofür wir die Energie brauchen - Wissenschaftliches Hochleistungsrechnen am Beispiel: Technische Strömungssimulation



- **Multiphysik:** steigende Komplexität von Problemen und Methoden
- **Modelle sind vielseitig:** Anwendung in Maschinenbau, Medizin, Fahrzeugen, ...
- **Diskrete Kontinuumsmechanik:** Vorhersage physikalischer Quantitäten an sehr vielen Punkten im Raum, z.B.: Druck, Temperatur, ..., und: viele Zeitpunkte
- **Hoher Ressourceneinsatz:** Speicher, Zeit, Energie -> Rechenzentrum = paralleler Großrechner

Computer allein werden bereits bis zur Mitte des Jahrhunderts mehr Energie benötigen, als produziert werden kann



Rechenleistung und elektrische Leistungsaufnahme im Supercomputerbau, damals

(ICARUS Designjahr **2014**)

Green500 Platzierung	Architektur	Top500 Platzierung	Performance / P [GFlop/s / W]	P [MW]	
1	Xeon + FirePro	168	5,2	0,057	<ul style="list-style-type: none"> ▫ Kein Top500 Top-Scorer (Top500#1:P=18) ▫ Xeon CPU: konventionell ▫ GPUs (Fire,Kxx): Beschleuniger ▫ PEZY CPU: unkonventionell
2	Xeon + PEZY	369	4,9	0,037	
3	Xeon + K20x	392	4,4	0,035	
4	Xeon + K40m	361	4,0	0,044	

www.top500.org/green500/list/2014/11

Rechenleistung und elektrische Leistungsaufnahme im Supercomputerbau

(heute)

Green500 Platzierung	Architektur	Top500 Platzierung	Performance / P [GFlop/s / W]	P [MW]	
1	Xeon + P100	28	9,5	0,350	<ul style="list-style-type: none"> ▫ Perf./P: x1,5! ▫ Sunway: auch unkonventionell und Top500#1! ▫ Powerdown Top500#1: x0,8
2	Xeon + P100	8	7,5	1,310	
3	Xeon + PEZY	116	6,7	0,150	
4	Sunway	1	6,0	15,370	

www.top500.org/green500/list/2016/11

“

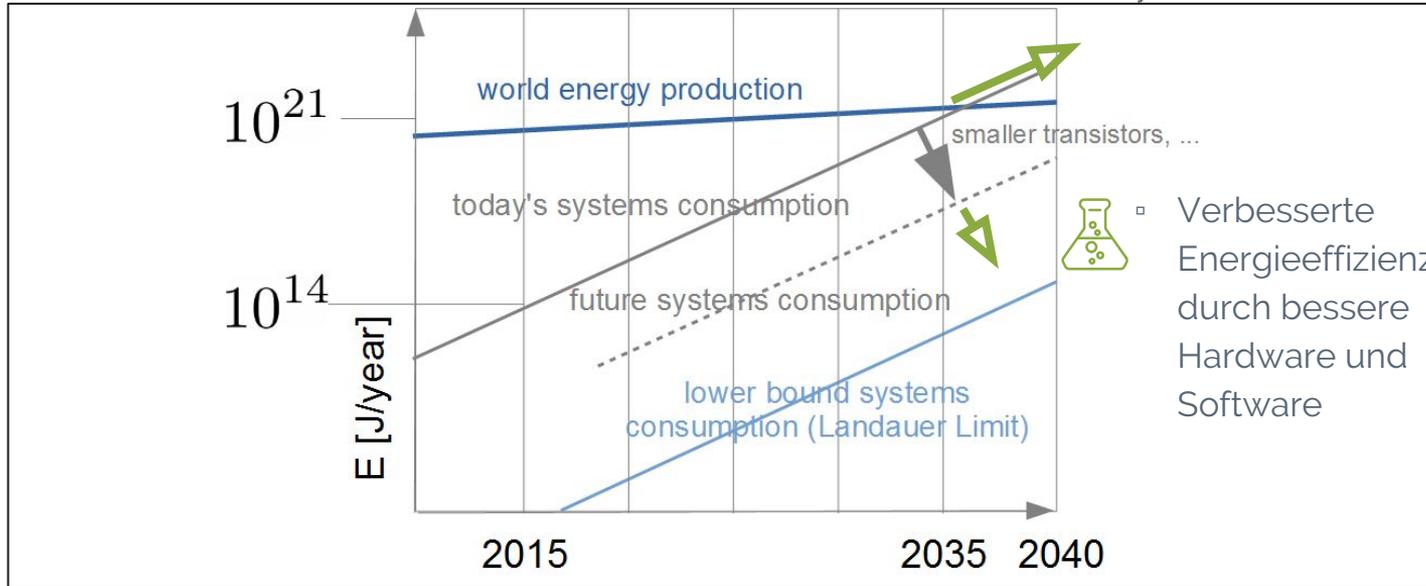
**Conventional approaches are running into physical limits.
Reducing the 'energy cost' of managing data on-chip
requires coordinated research in new materials, devices,
and architectures.**

- the Semiconductor Industry Association (SIA), 2015

Was können wir im Scientific Computing tun?



- Integrierte erneuerbare Energiequelle für große Systeme



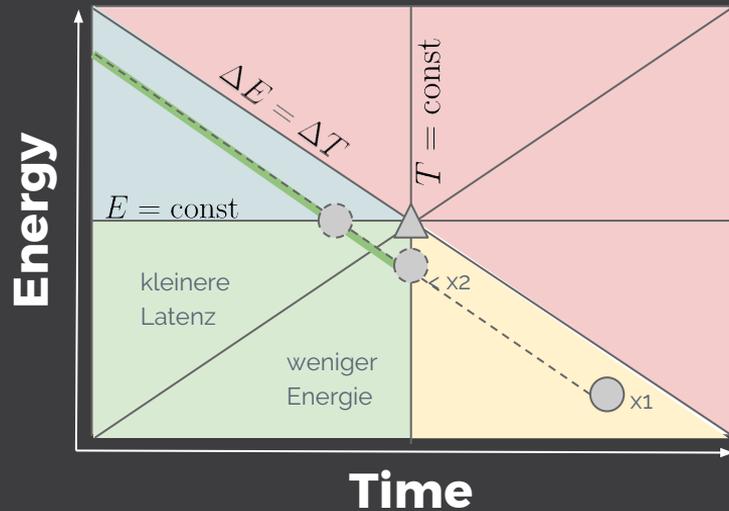
- Verbesserte Energieeffizienz durch bessere Hardware und Software



2

Hardware, Numerik und Energieeffizienz

Wie wir Mobile Computing nutzen können

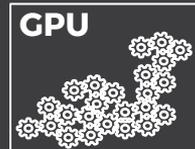
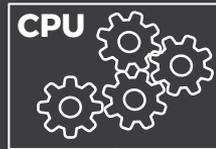
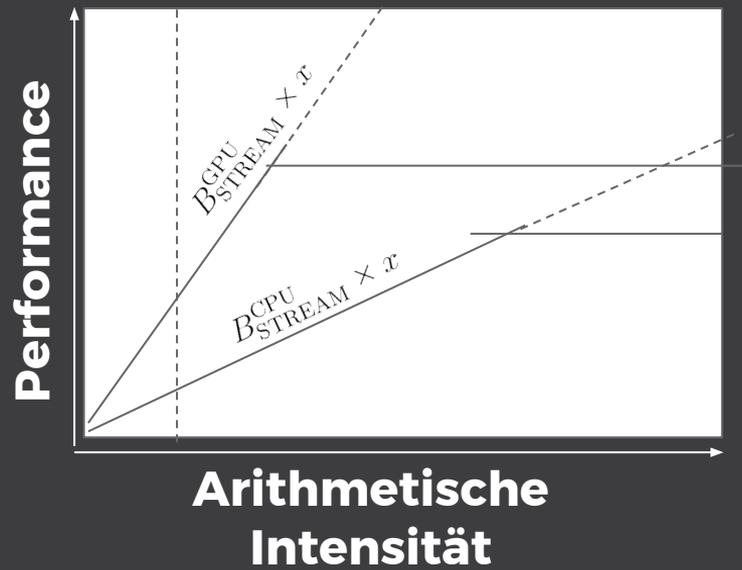


Hardware und Numerik relativ zu Performance, Power, Energy

- **Theoretische Peak Performance** ist gegeben durch die Mikroarchitektur der Hardware
- **Maximal erreichbare Performance** ist eine Funktion von Hardware und Task
- **Tatsächliche Performance** ist eine Funktion von Hardware und Implementierung (Numerische Effizienz und Parallelisierung), **Rechenzeit** ist invers dazu
- **Energy** (to solution) ist die Gesamtenergie für eine Ausführung des Codes und ergibt sich durch Integration der Leistungsaufnahme (**Power**) in der Zeit
- **Speed-/Perf.- bzw. Powerdown bzw. -up** sind die Faktoren zwischen zwei Hardwarearchitekturen bezüglich numerischer bzw. elektrischer Leistung bei gleichem Task

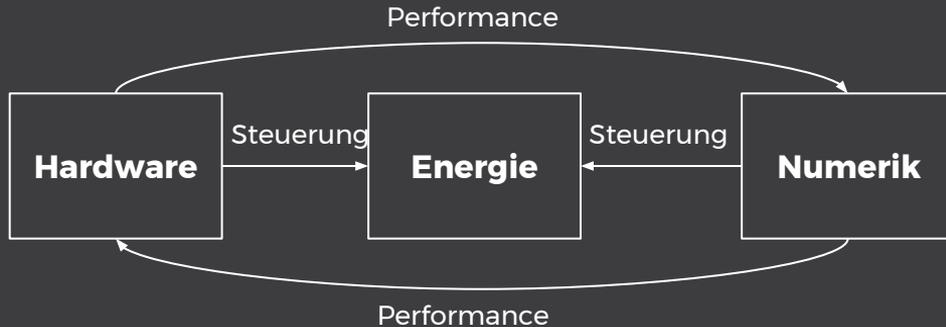
Damit dieselbe Simulation auf einer alternativen Hardware energieeffizienter ist, muss (mindestens) gelten:
Powerdown > Speeddown.

Hardware und Numerik relativ zu Performance, Power, Energy



- **Maximal erreichbare Performance** ist eine Funktion von Hardware und Task
- **Hardwarearchitekturen** sind sehr unterschiedlich
- **Tasks** haben unterschiedliche Anforderungen an die Hardware
- Wir (Strömungssimulation!) brauchen oft hohe on-chip **Speicherbandbreite** und **feingranulare Parallelität**

Energieeffizienz und Hardware-orientierte Numerik



- Performance und **Effizienz** sind nicht dasselbe
- **Methoden** können teurer sein als andere aber eine höhere **numerische Effizienz** haben, wodurch die Rechenzeit insgesamt sinkt
- Mehr Kerne oder eine höhere Kerntaktfrequenz in einem Mehrkernsystem zu nutzen kann eine Simulation beschleunigen, aber die **Energieeffizienz** kann sinken
- Sehr komplexe Dinge gehen in einem Core vor (DFS/DVS, pipelining, out-of-order execution, caches). Bereits das Erreichen einer guten **Hardwareeffizienz** auf einzelnen Cores ist schwierig aber nötig. Erst dann macht **Parallelisieren** Sinn.

Mobilprozessor Tegra K1

Entwicklungshistorie: Muss von (kleiner) Batterie versorgt werden

32 Bit Architektur,
4 x Cortex-A15 CPU,
Programmierbare Kepler GPU,
LPDDR3,
Hohe SP Performance,
Energieeffizienteste Plattform
ihrer Zeit am Consumer-Markt,
Inzwischen: Tegra X1, Tegra X2,

$$\frac{\text{Perf}_{\text{peak}}^{\text{TegraK1}}}{P_{\text{peak}}^{\text{TegraK1}}} = \frac{35\text{GFlop/s}}{W}$$

$$\frac{\text{Perf}_{\text{peak}}^{\text{Ref}}}{P_{\text{peak}}^{\text{Ref}}} = \frac{11\text{GFlop/s}}{W}$$

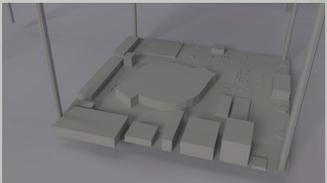
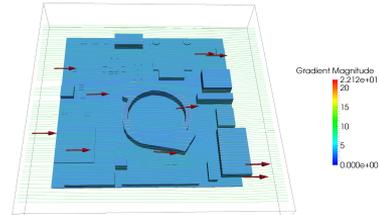
Ref.: Xeon + K40

...

Jetson TK1 - Der Bastelrechner

Alles, was man in einem Rechenknoten braucht

Tegra K1 Carrier Board,
P = 10 - 15W mit Lüfter,
GiBit Ethernet,
Viel I/O, z.B.: Seriell, USB, SATA
Linux, CUDA
Inzwischen: Jetson TX1, Jetson TX2



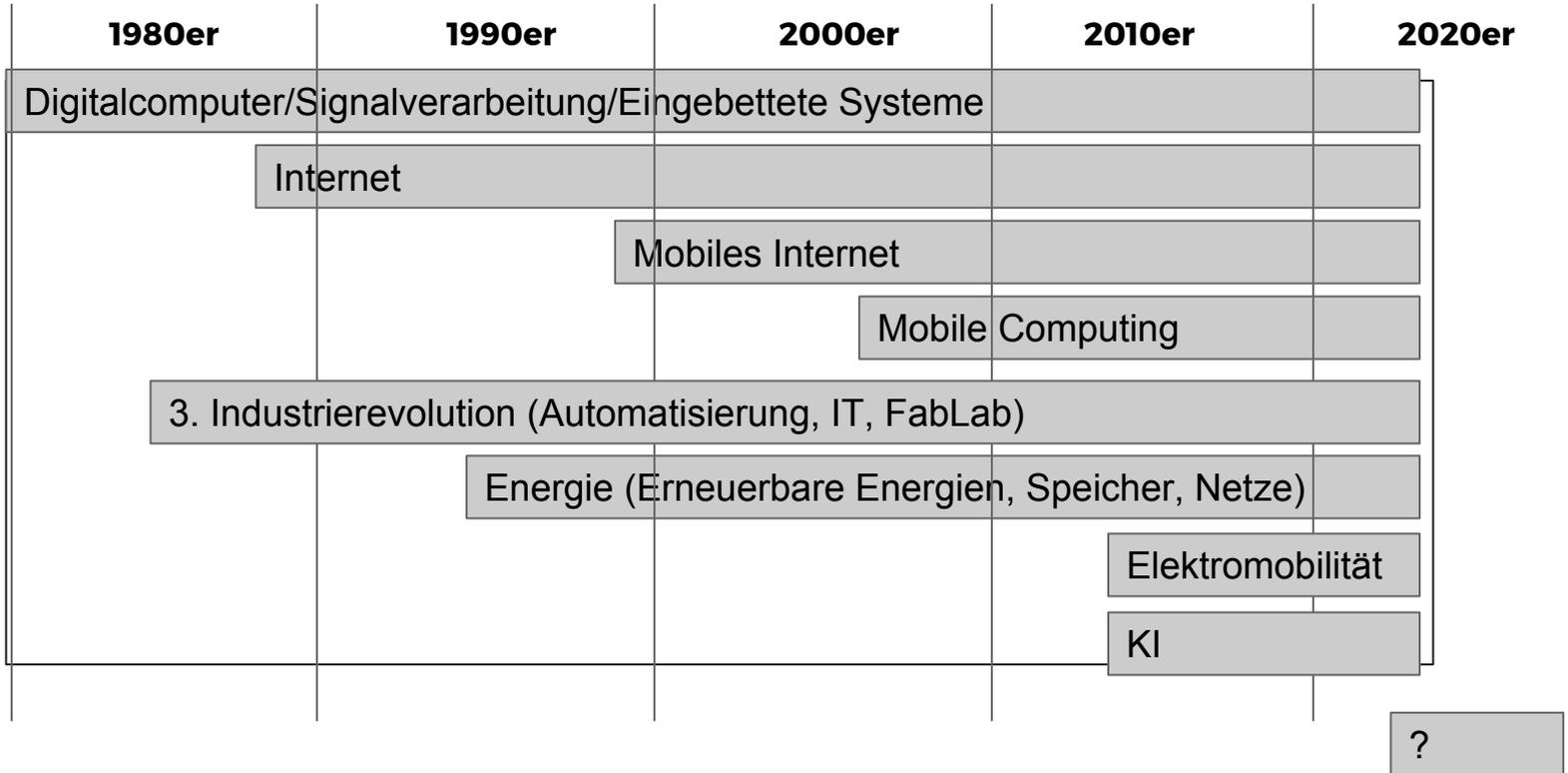


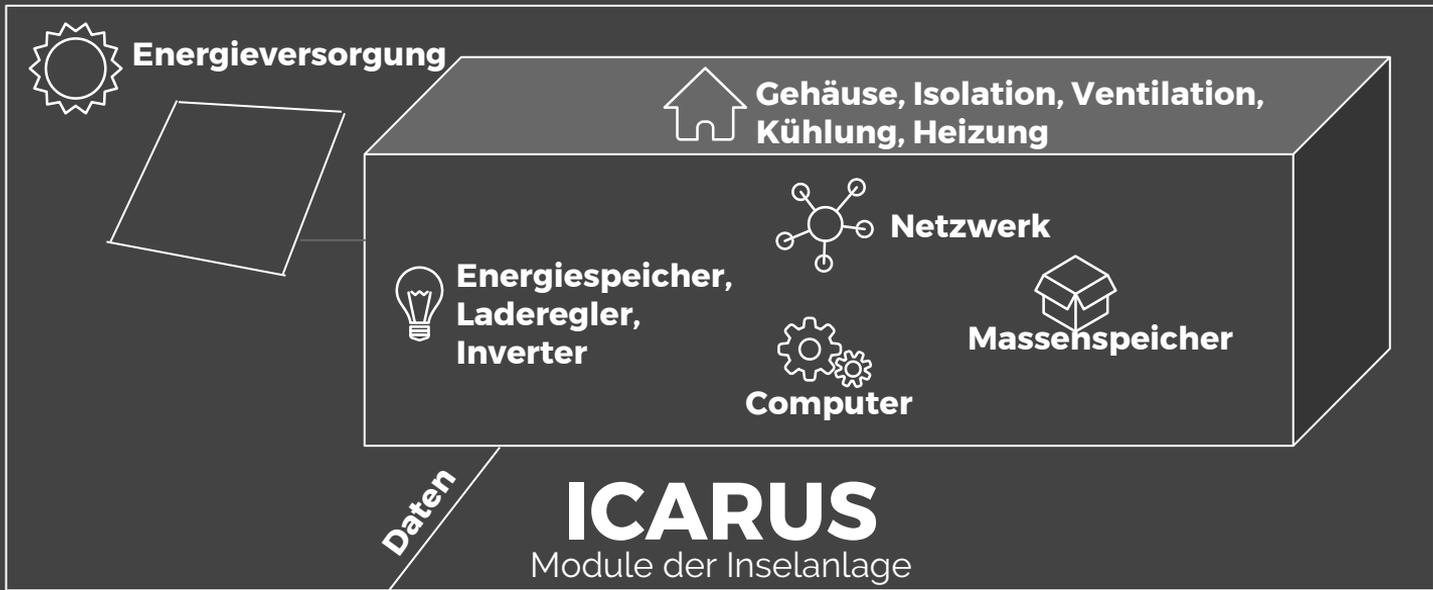
3

Systemintegration mit Energiewendetechnologien

Technologiesynthese für ein Rechenzentrum der Zukunft

Technologierevolutionen und Systemintegration







Energieversorgung & -speicher

Herausforderungen: Wetter, benötigte Fläche, Nachtbetrieb

7,5 kWp Photovoltaik, Freifläche
2,5m x 16m,
8kWh Li-Ionen Akku,
2+1 Inverter, Laderegler



Zielvorgabe: Volllastbetrieb
ganztäglich im Sommer +
Batterieladung; in der Nacht,
Winter: am Leben bleiben



Gehäuse, Kühlung, Heizung

Herausforderungen: Isolation, benötigte Fläche, Lüftung, Brandschutz

Umgebauter Seefrachtcontainer,
90mm Isolationswände, brandgeschützt,
Stahlsicherheitstüren,
Ventilation mit eigener Energieversorgung,
Ventilation heizbar am Tag,
Zusätzliche Heizung für Batterie bei Nacht



Cluster, Netzwerk, Massenspeicher

Herausforderungen: Raumausnutzung, Wärmenestvermeidung

60 x NVIDIA Jetson TK1, 240 ARM cores, 60 GPUs,
120 GB RAM (LPDDR3),
3 PDUs, Sensoren,
3 + 1 Low Power Switches (Netgear)
Leistungsaufnahme (peak): < 1kW
Theoretische Peak Performance: 20 TFlop/s mit gemischter Genauigkeit

Energieeinsparung beim Massenspeichersystem mit Bastelrechnern

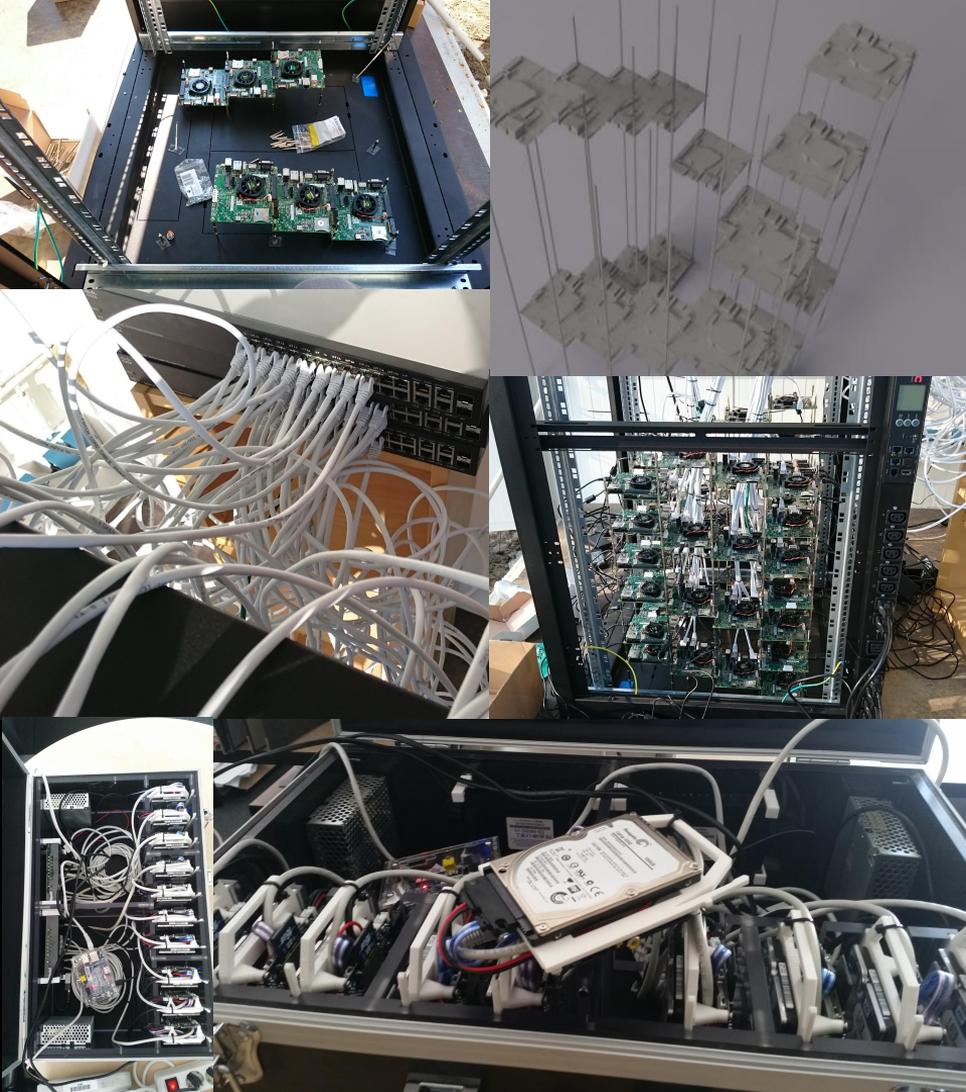
Voll portabel dank BananaPi,
10 x 1 TByte SSDs,
Redundant, z.B.: 5 TByte nutzbar



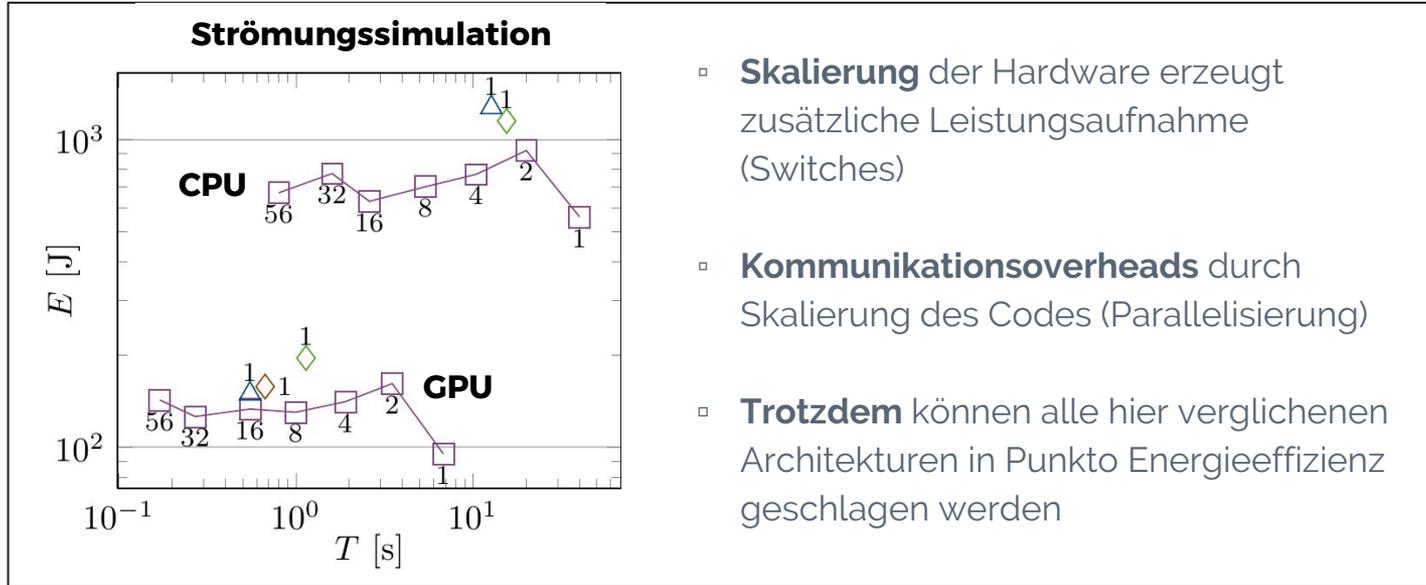
MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

Maximal gemessene Zugriffsgeschwindigkeit:
90 MByte/s (zum Vergleich: 140 bei Standardsystem) -> **Speeddown = 1,6**

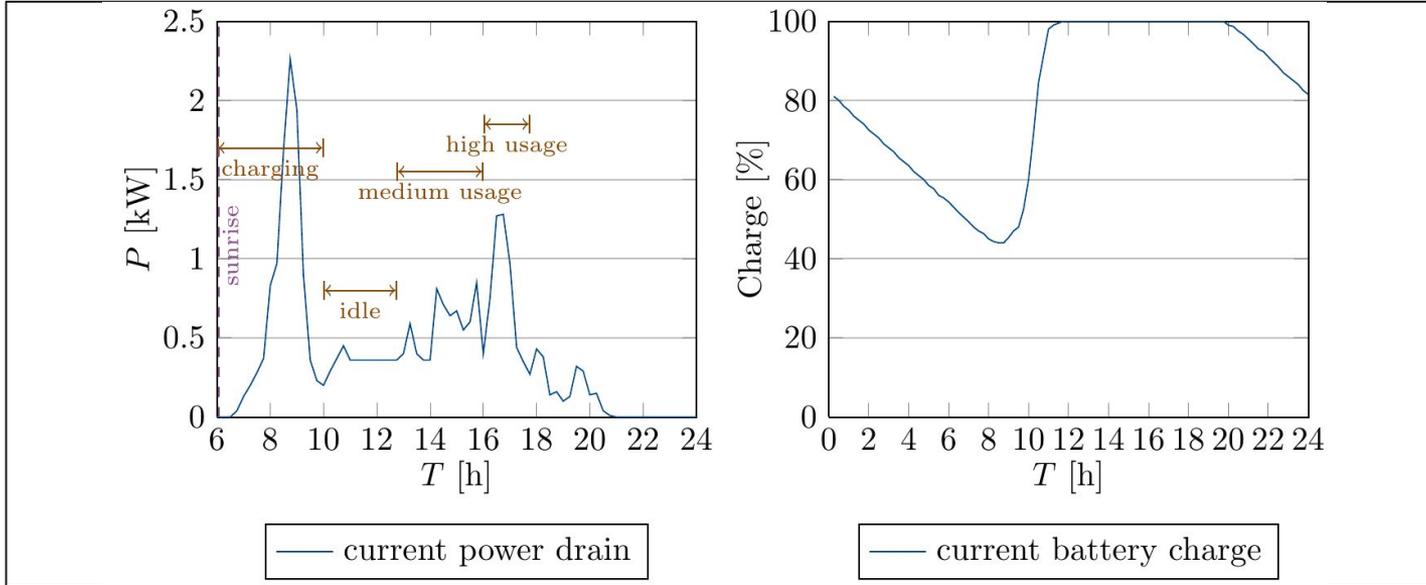
Durchschnittliche Leistungsaufnahme: 30W (zum Vergleich: 500W bei
Standardsystem) -> **Powerdown = 16,8**



Ergebnisse Energieeffizientes Wissenschaftliches Rechnen



Ergebnisse Solarenergie



Verlässlichkeit

Trotz starker Belastungen (Temperatur):
Keine größeren Ausfälle seit Fertigstellung März
2016

Klima im Container

An heißen Tagen:
Max 33 °C ambient,
Max 68 °C Jetson boards,
Luftfeuchte unter Kontrolle

An kalten Tagen: Batterie
warmhalten schwierig

ICARUS
Im Einsatz



Danke!



ICARUS @ PACO17 Magdeburg: <https://www.mpi-magdeburg.mpg.de/csc/events/paco17>