Resource-aware Research on Universe and Matter: Call-to-Action in Digital Transformation

Ben Brüers[a]

[a]Deutsches Elektronen Synchrotron DESY

Sustainable HEP 2024

arXiv:2311.01169

Monday, 10.06.2024, 16:10

HELMHOLTZ SPITZENFORSCHUNG FÜR GROSSE HERAUSFORDERUNGEN

DESY.

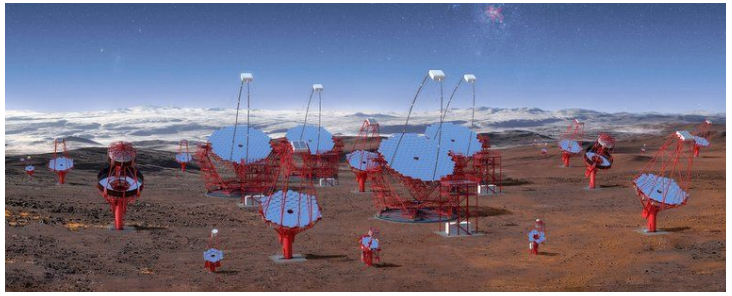
What is Research on Universe and Matter?

- Research on Universe & Matter = research program by German ministry
- Supports large scale research facilities and infrastructures
    - e.g. LHC+experiments, XFEL, PETRA, BESSY, FAIR, CTA, VLT
- Spans Particle Physics, Nuclear Physics, Astro(particle) Physics, Material Science, Condensed Matter Physics, Biophysics, Accelerator Physics

https://www.swr.de/wissen/1000-antworten/wie-gelaehrlich-ist-der-lhc-teilchenbeschleuniger-am-cern-in-genf-100.html

https://www.xfel.eu/news_and_events/news/index_eng.html

https://www.eso.org/public/germany/announcements/ann230009/

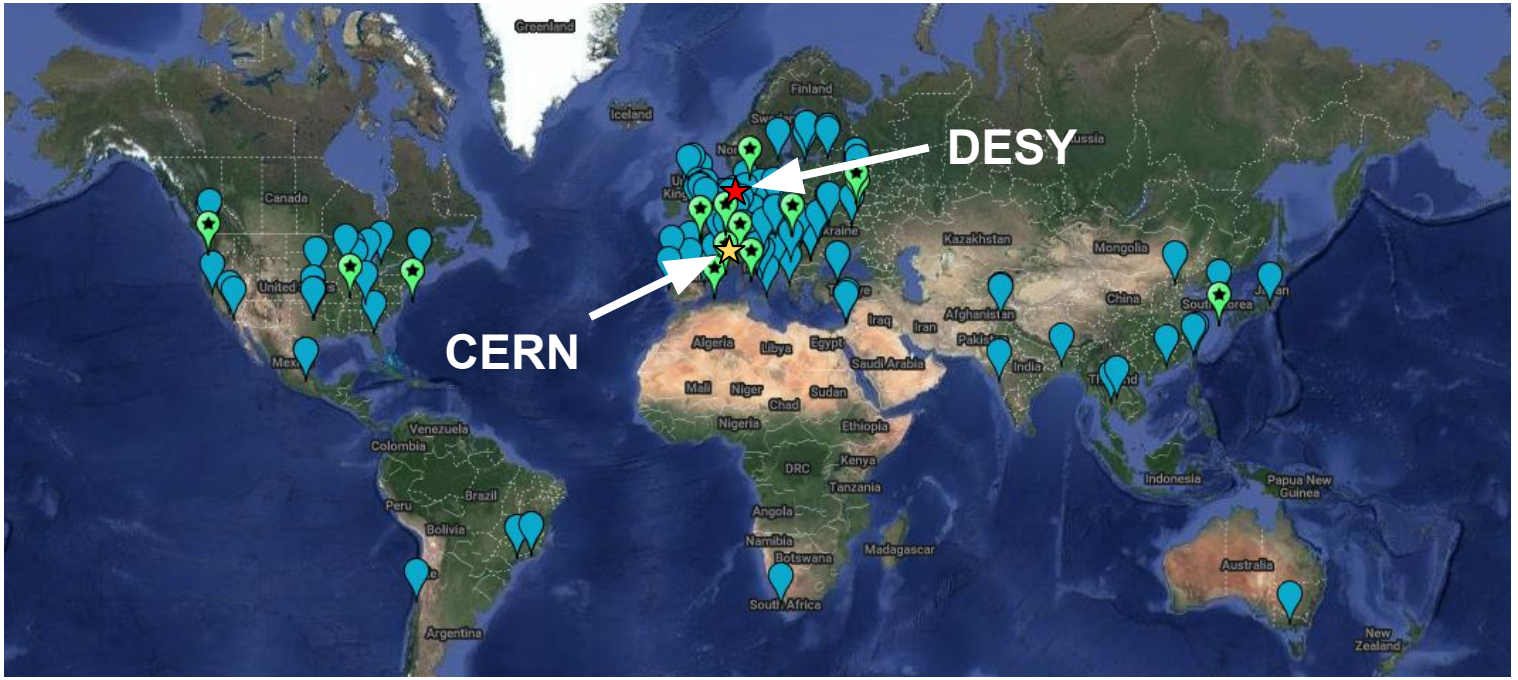


https://www.mpi-hd.mpg.de/mpi/de/forschungsbeteiligungen-und-projekte/nichtthermischeastroteilchenprojekte/cta

Example: Particle Physics and Computing

- Many of our Particle Physics facilities…
  - …produce large amounts of data → needed to identify rare events
  - …use complex simulations to compare predictions to measured data

→ **Large scale computing facilities needed! Use cloud-like infrastructure**



https://wlcg.web.cern.ch/using-wlcg/monitoring-visualisation

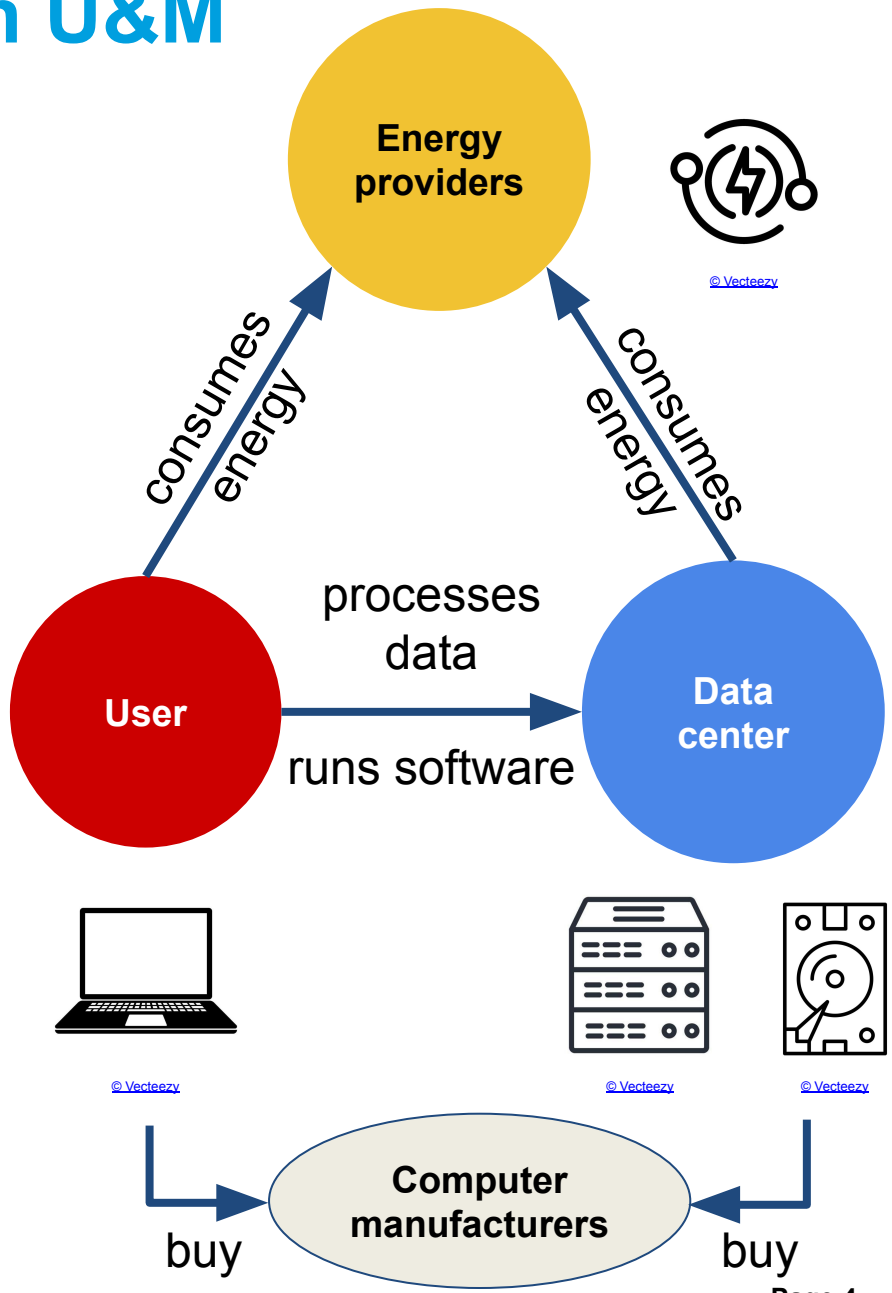
LHC computing center sites (=WLCG)

Computing resources in U&M

- Large-scale computing centers process & store data in U&M
  → **non-zero CO2e footprint**

- Workshop in May/June 2023 by U&M computing community to reduce CO2e
  - Main incentives: **reduce, reuse, recycle, avoid!**
  - Follow-up workshop to monitor progress

Workshop paper:
arXiv:2311.01169



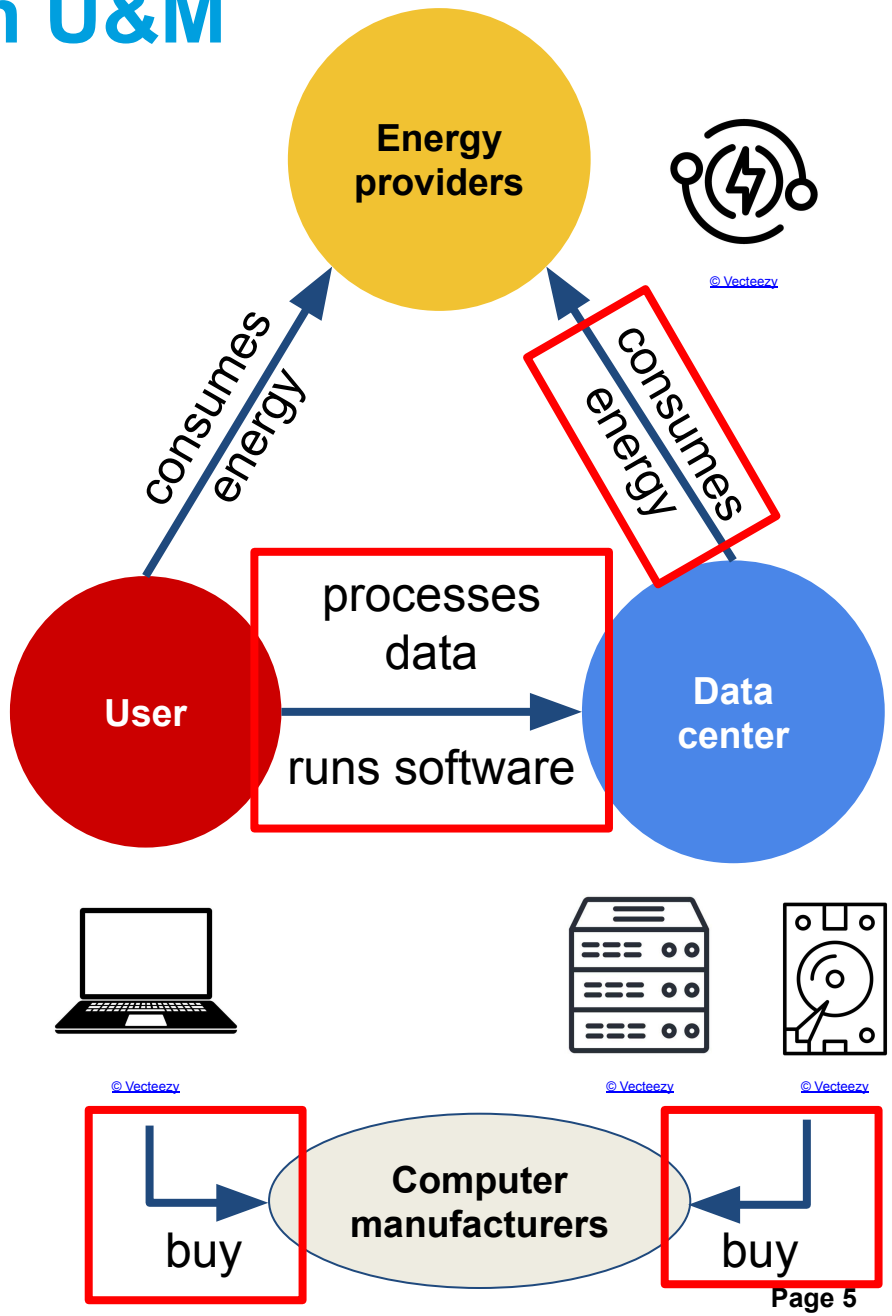
Computing resources in U&M

- Large-scale computing centers process & store data in U&M
  → **non-zero CO2e footprint**

- Workshop in May/June 2023 by U&M computing community to reduce CO2e
  - Main incentives: **reduce, reuse, recycle, avoid!**
  - Follow-up workshop to monitor progress

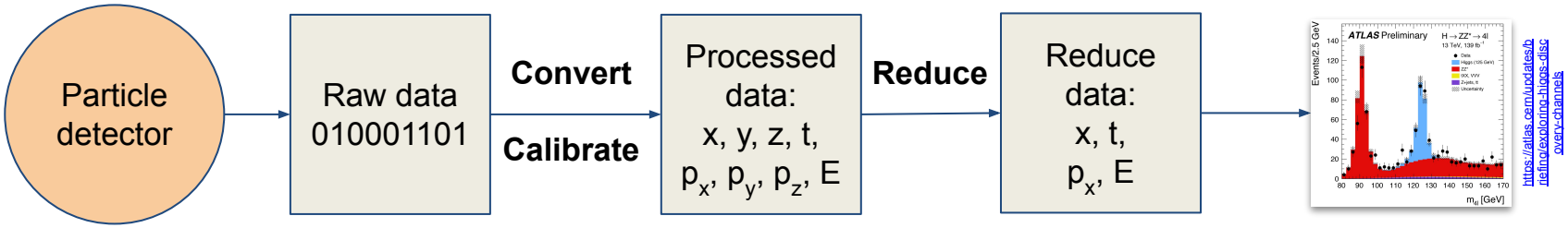
Workshop paper:
arXiv:2311.01169

Will discuss these



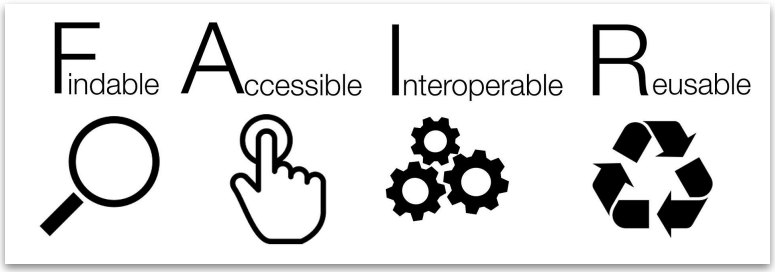
Process your data smartly!

- Detectors output bit-stream → transformation to be human readable
- Data accessed multiple times in data analysis
  → additional (temporary) transformation efficient at the cost of extra storage
- **Trade off between transforming / storage → "smart" transformations**

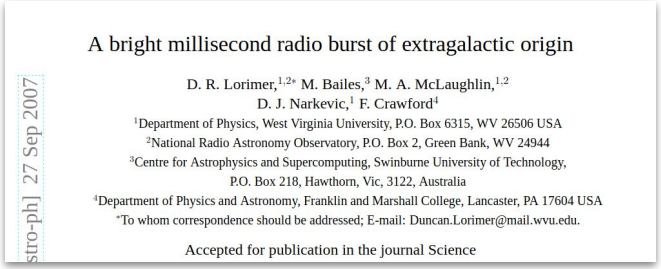


https://atlas.cern/updates/briefing/exploring-higgs-discovery-channels

- Reusing existing data is efficient → has led to several surprises in the past
- **Smart transformations for easy, long-term access of data → FAIR principles**



https://en.wikipedia.org/wiki/FAIR_data#/media/File:FAIR_data_principles.jpg


A bright millisecond radio burst of extragalactic origin

D. R. Lorimer,[1,2*] M. Bailes,[3] M. A. McLaughlin,[1,2]
D. J. Narkevic,[1] F. Crawford[4]
[1]Department of Physics, West Virginia University, P.O. Box 6315, WV 26506 USA
[2]National Radio Astronomy Observatory, P.O. Box 2, Green Bank, WV 24944
[3]Centre for Astrophysics and Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, Vic, 3122, Australia
[4]Department of Physics and Astronomy, Franklin and Marshall College, Lancaster, PA 17604 USA
*To whom correspondence should be addressed; E-mail: Duncan.Lorimer@mail.wvu.edu.

Accepted for publication in the journal Science

arXiv:0709.4301

Use efficient software!

Need to process lots of data
→ write efficient software!

- Re-use existing software
  - more users = better
  - **requires maintainability, documentation, versioning**
  - use existing libraries/algorithms
- Test / benchmark
- Adjust to hardware (e.g. ARM)
- **Use artificial intelligence, but be critical**

https://upload.wikimedia.org/wikipedia/commons/thumb/e/e0/Git-logo.svg/1280px-Git-logo.svg.png

https://en.m.wikipedia.org/wiki/File:Boost.png

https://upload.wikimedia.org/wikipedia/commons/thumb/3/31/NumPy_logo_2020.svg/1280px-NumPy_logo_2020.svg.png

nature › scientific data › articles › article

Article | Open access | Published: 14 October 2022

Introducing the FAIR Principles for research software

Michelle Barker, Neil P. Chue Hong, Daniel S. Katz, Anna-Lena Lamprecht, Carlos Martinez-Ortiz, Fotis Psomopoulos, Jennifer Harrow, Leyla Jael Castro, Morane Gruenpeter, Paula Andrea Martinez & Tom Honeyman

Scientific Data 9, Article number: 622 (2022) | Cite this article

19k Accesses | 55 Citations | 232 Altmetric | Metrics

https://www.nature.com/articles/s41597-022-01710-x

Use your hardware as long as possible!

- **Also production matters for CO2e!**
  - 75% for a PC
  - 50% for a server
- **Use hardware as long as possible!**
  - Replace/fix failing parts
- But: new hardware may be more efficient! → Trade-off
- Construct “old hardware” computing centers running only on green energy?

13-inch MacBook Pro life cycle carbon emissions

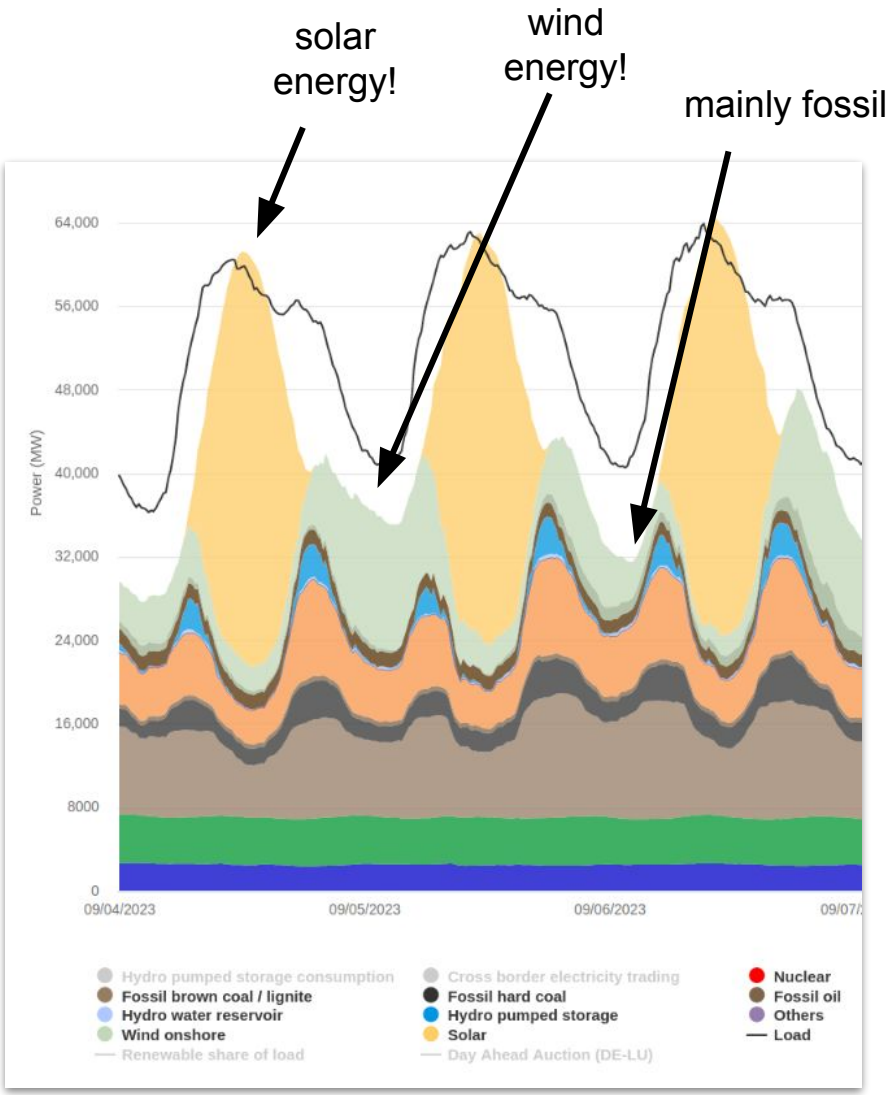
73%	Production
7%	Transport
19%	Use
<1%	End-of-life processing

https://www.apple.com/environment/pdf/products/notebooks/13-inch_MacBookPro_PER_Nov2020.pdf

https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/Full_LCA_Dell_R740.pdf

Monitor & optimise computing centers!

- **Computing centers should monitor energy usage**
  - Breakdown to understand where to save
- **Use only renewable energy?**
- Avoid energy transportation losses → new centers close to renewables
  - National German scientific computing center at north sea?
- **Challenge: green energy not constantly available**

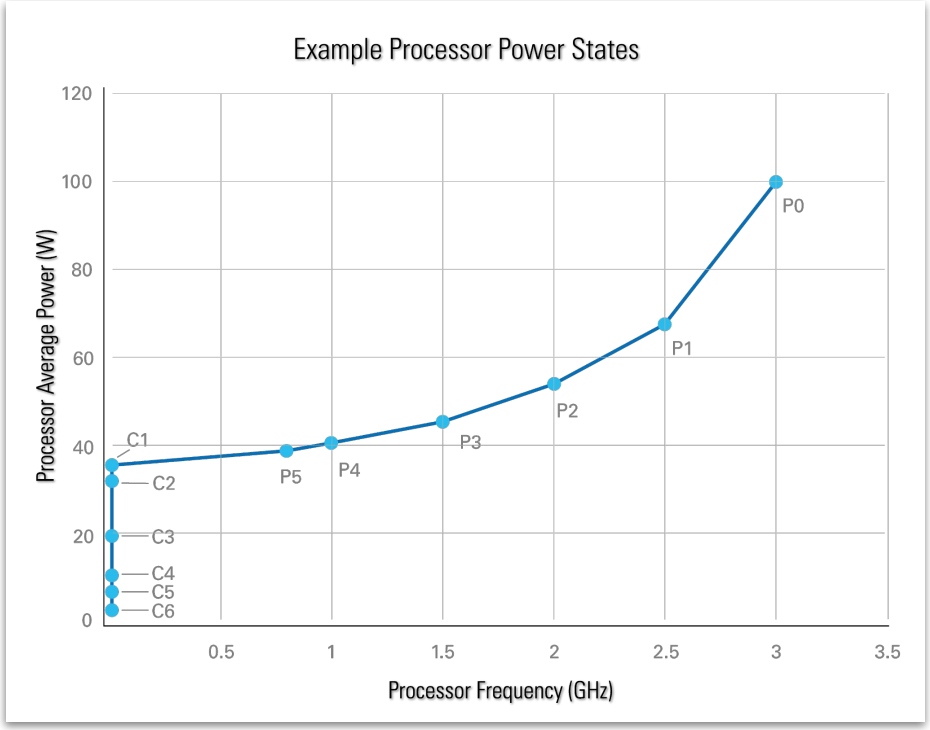


Rodney Walker:
https://indico.desy.de/event/37480/contributions/140510/attachments/82246/108365/Meinerzhagern_compOps(2).pdf

Tackling varying energy supply

- Turn computers off → lifetime issues
  - **Energy storage needed**
- Killing jobs no option
  - Checkpoint / “freeze” jobs
  - **Reduce CPU clock frequency**
    - Tested:
      → similar performance
      → but jobs take longer
- **Longer jobs no problem if running O(days) anyway**

https://www.wsj.com/business/energy-oil/giant-batteries-helped-the-u-s-power-grid-eke-through-summer-a68425fd



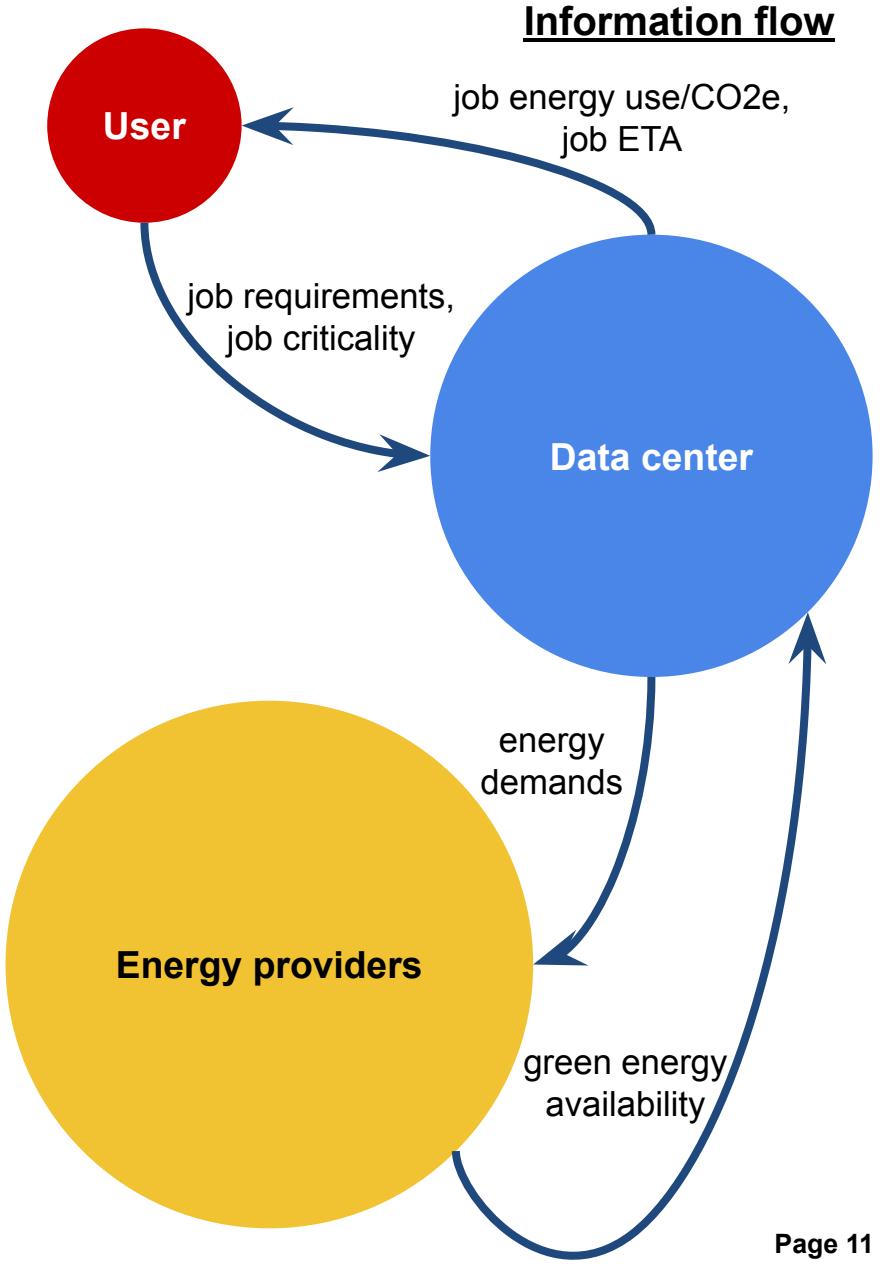
Rodney Walker:
https://indico.desy.de/event/37480/contributions/140510/attachments/82246/108365/Meinerzhagern_compOps(2).pdf

Frequency [GHz]	Power [W]	Calculations / W [HS06]	Calculations / W / nominal
1.5	286	3.79	98%
2.15	330	4.32	111%
2.85	524	3.88	100%

T2 AMD, Thomas Hartmann, DESY

We need new information flows & tools!

- **Reliable green energy operation**
  → new information flows & tools
- **Need new tools!**
  - information to/from users
  - information exchange with power provider
  - data center job/energy control
- **Already today: report CO2e** of jobs/users/data-centers
  → help make software efficient
  - e.g. NAF @DESY

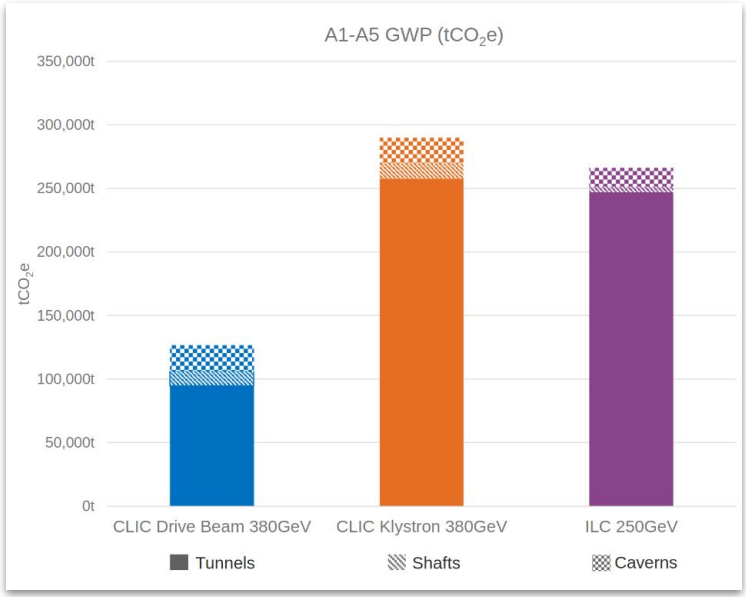


Our culture has to change!

- **Raise awareness & train people**
  - Efficient software
  - Avoid unnecessary repetitions
  - Document workflows
- Consider CO2e budgets etc. (“carrot-and-stick”)

- **Assess & reduce CO2e footprint of current projects**
- **Consider sustainability in all future projects**

https://indico.desy.de/event/40133/



https://edms.cern.ch/document/2917948/1

Funding is adjusting!!

- E.g. sustainability strategies from German government / ministry → “Research and sustainability”
- **Now have explicit funding opportunities for sustainability in U&M research**
- **Need solid structural basis to enable future sustainable computing/research**

https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/7/31638_Forschung_fuer_Nachhaltigkeit.pdf?__blob=publicationFile&v=6

CLIMATE CHANGE WILL NOT WAIT FOR US TO FINISH OUR RESEARCH.

LET'S TAKE ACTION NOW!

Let's bundle our efforts!
Get in touch at info@erumdatahub.de!

Workshop paper:
arXiv:2311.01169

Thank you

Let’s bundle our efforts!
Get in touch at info@erumdatahub.de!

Workshop paper:
arXiv:2311.01169

Contact

DESY.
Deutsches Elektronen-Synchrotron

www.desy.de

Ben Brüers
ATLAS group (Zeuthen)
ben.brueers@desy.de
+49 33762 7-7640