Sustainability in HEP Computing

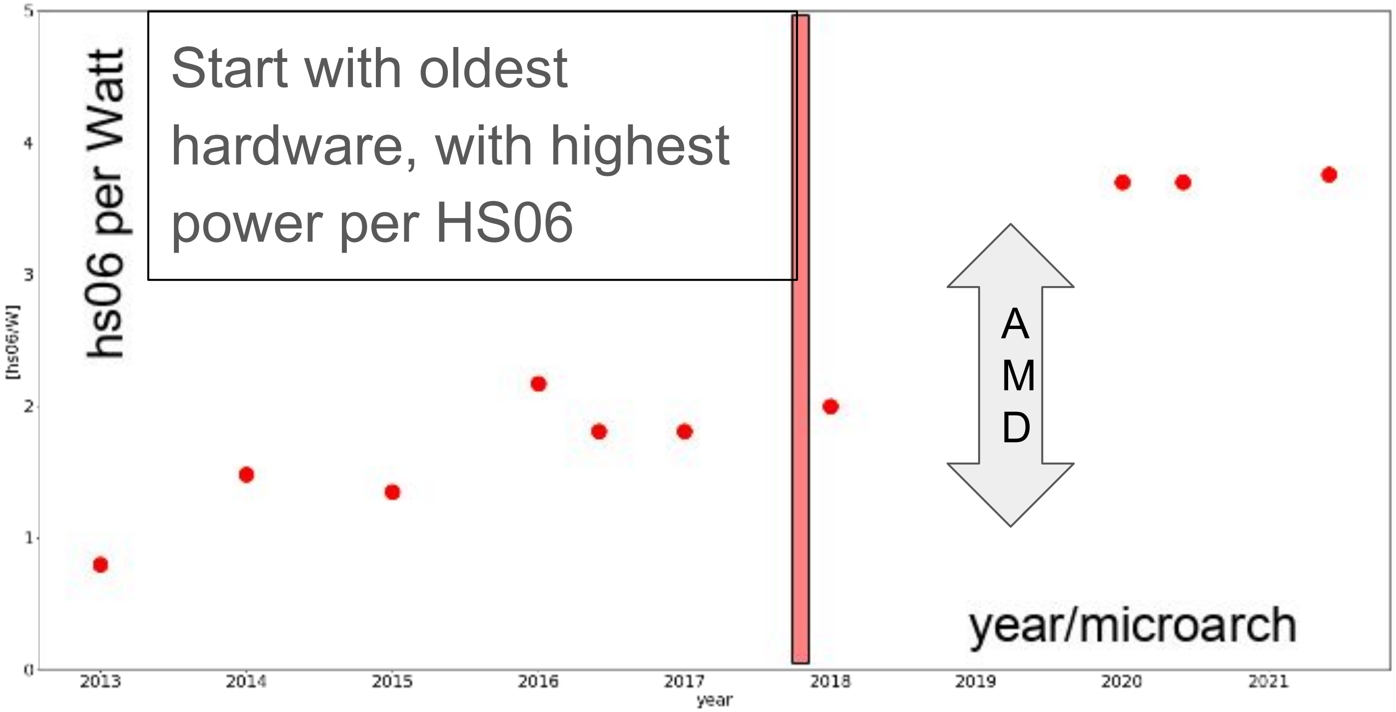
Rod Walker, LMU, Munich

Hot topic due to energy crisis in Europe, particularly Germany.
Stays hot due to net-zero need to dramatically reduce fossil fuel dependence, in favour of intermittent renewables.
HEP Computing can adapt since naturally distributed and flexible

Turn things off! Simple yet effective.

Initial voluntary reaction to national and EU requests.

Save money when electricity tariff increased, to fit in flat budget

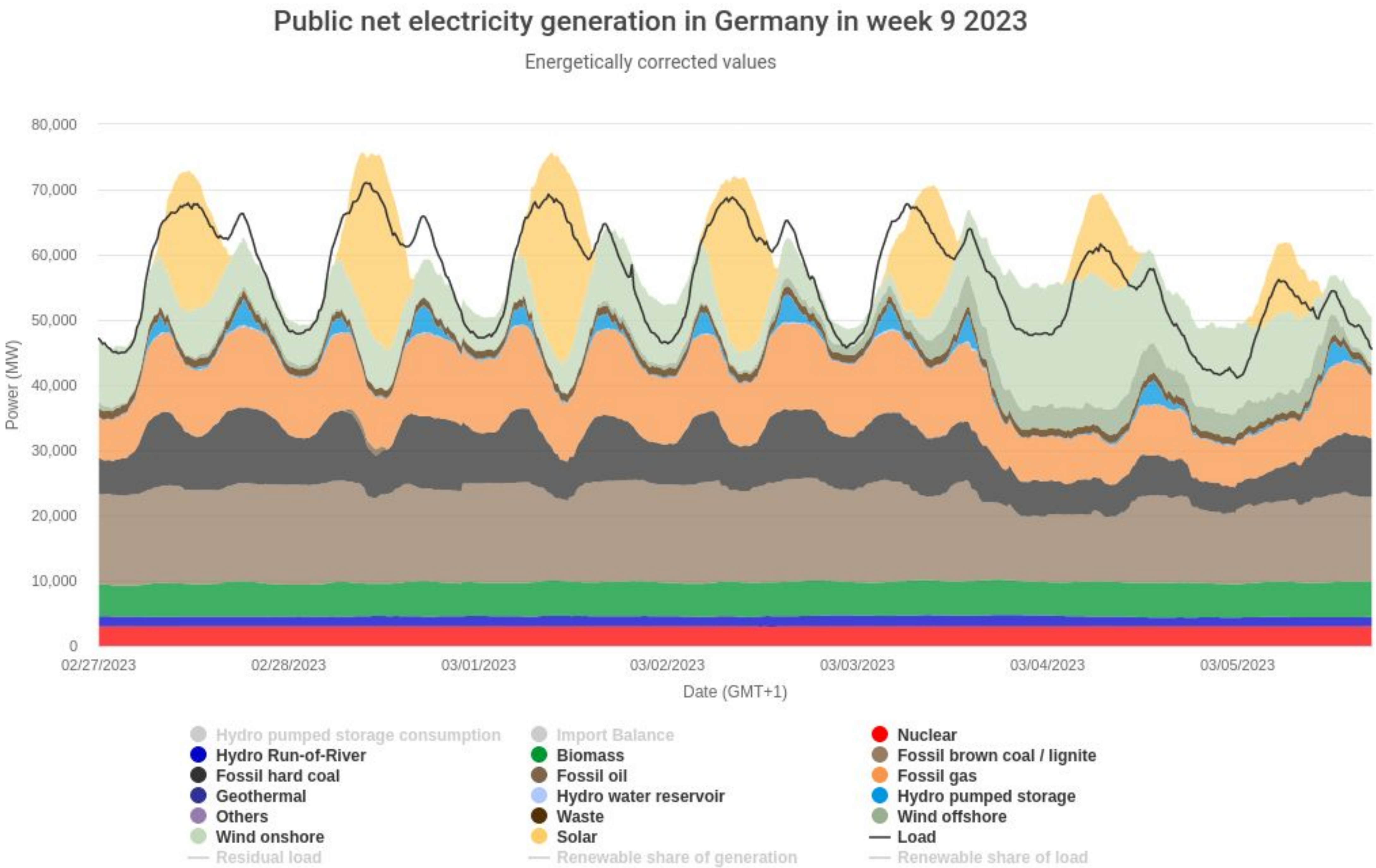


1. Retire older hardware.
2. Shutdown at low computing demand periods, e.g. analysis cluster during vacations, weekends & evenings.
Requires automatic **draining**, shutdown, power-on.
Draining depends on job lengths - wasted cycles.

But does flat reduction make sense?

- Demand peaks daily, less on the evening,weekend,holiday
- Solar production peaks midday, seasonal and not always, but forecast.
- Wind power is intermittent, but forecast.
- Difference between demand and renewables made up by fossil fuels
- Minimize the difference means dynamic **load-shedding**
- Modulate computing power usage typically twice per day
  - cannot switch off nodes at this frequency
    - draining time(no preemptable payloads) & hardware failure
- If renewables exceed demand, need load-shaping (increase useful demand)
  - run old hardware only when energy ‘free’, make up for load-shedding periods
  - regional excess often limited by distribution bottlenecks(location is important)

Proposal: Reduce CPU frequency to reduce compute power consumption by 50-60% at peaks in price. Run old hardware only when price is low.



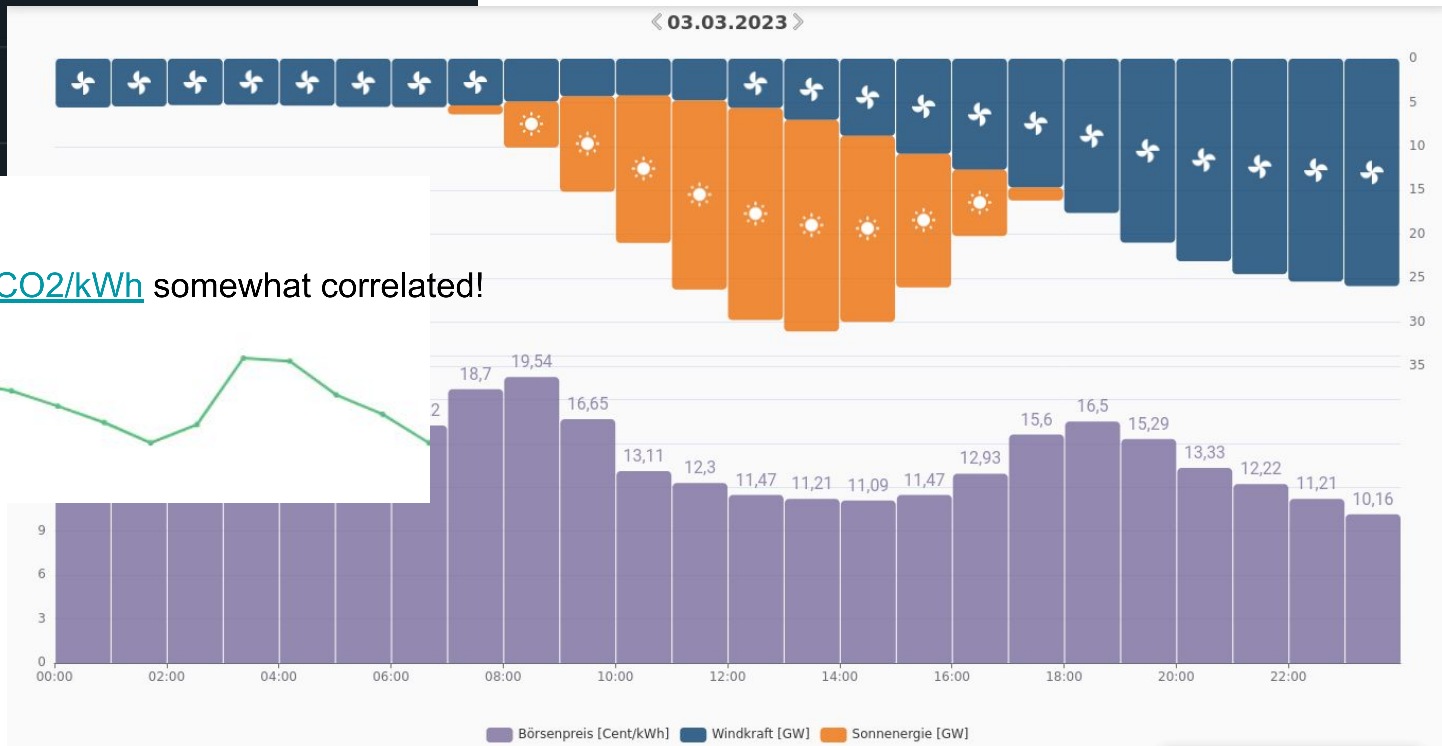
From
energy-charts.info

Variable electricity tariff

EEX day ahead price, plus base

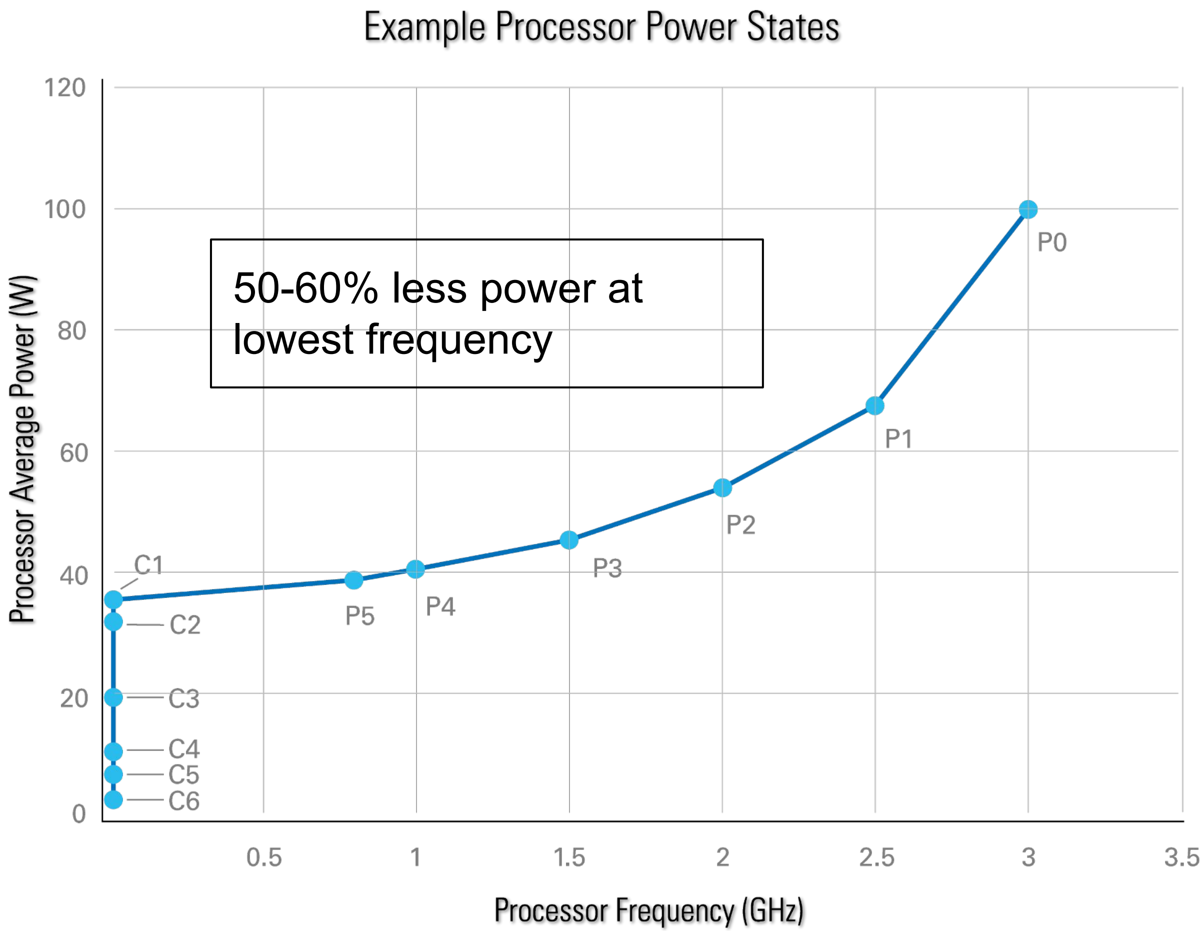
Tibber, Awattar but also business tarifs

gCO2/kWh somewhat correlated!



Save money by reducing power at peaks.
Save gCO2 when pricing politics catches up: higher carbon price/tonne

CPU frequency modulation fast & painless



$P \sim f V^2$

dynamic voltage and frequency scaling (**DVFS**)

voltage reduces with frequency

Useful work ~ frequency, but power falls faster than frequency

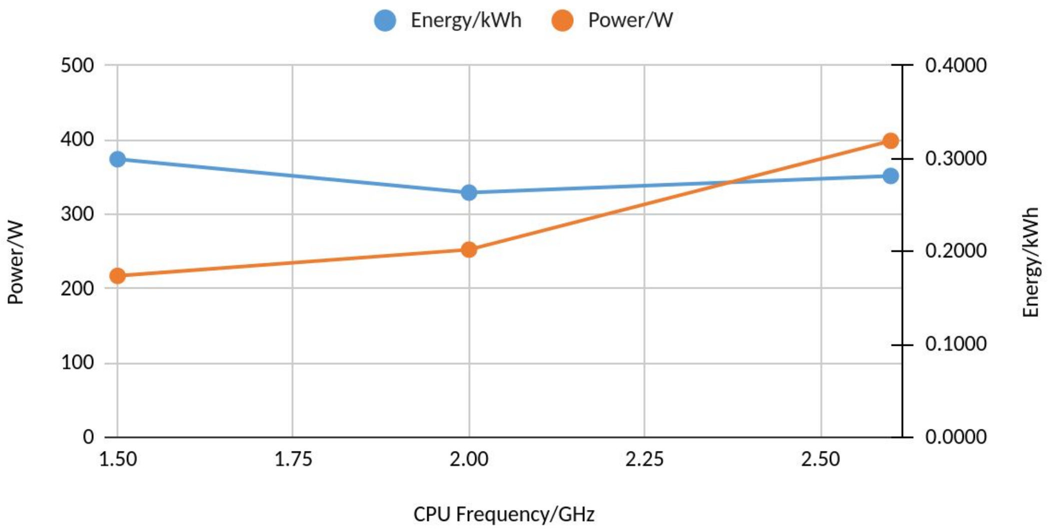
Offsets base/non-CPU node power consumption

No draining so can be frequent

Real-world measurements: HEP work vs total node power

dual-x86
E.Simili, Glasgow

1) Same work at different frequencies: 1000evt G4 sim

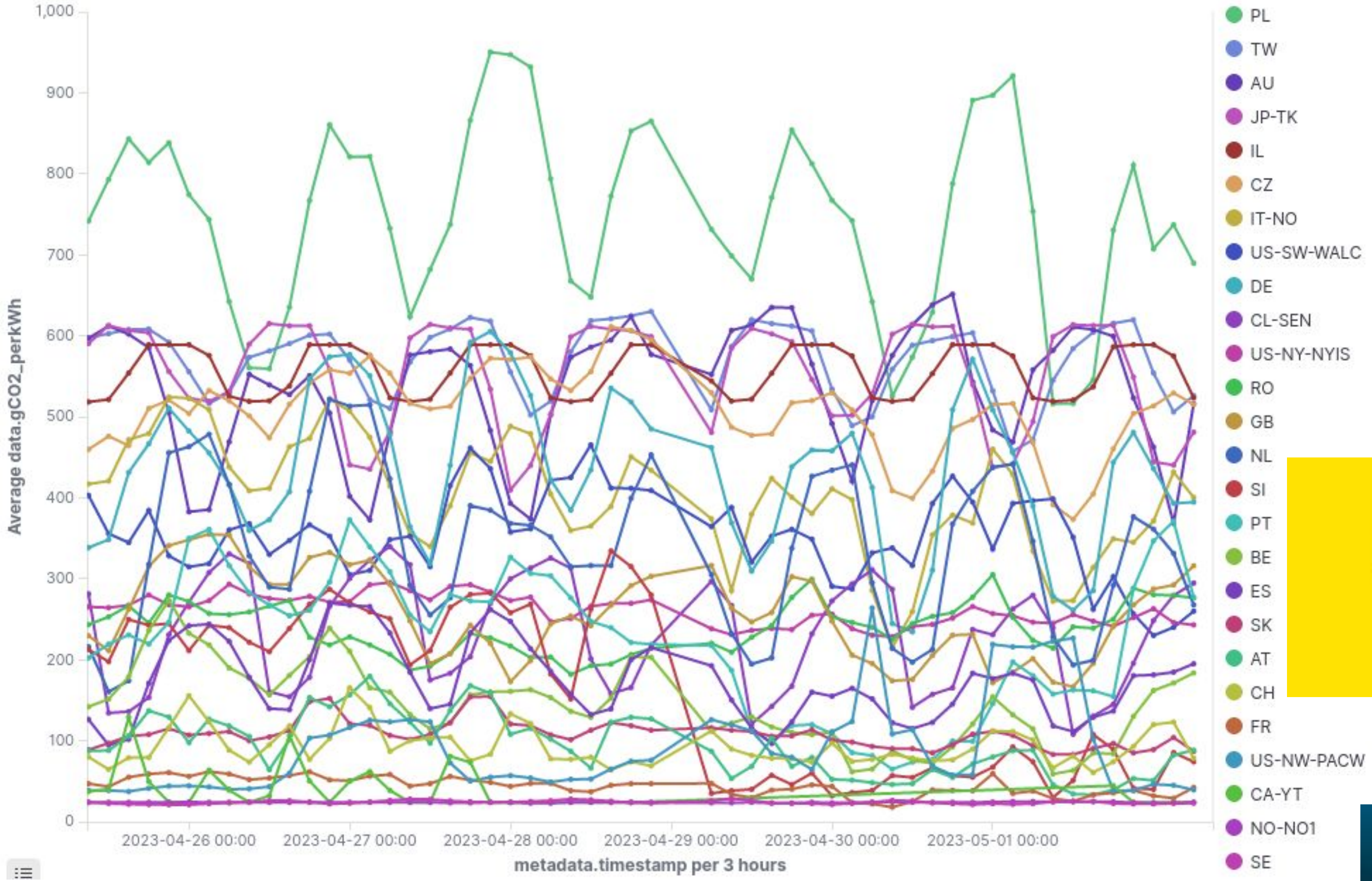


- **HEP work per kWh not significantly less at lowest frequency**
  - Glasgow 6% & DESY 2%
- Middle frequency best for both!
  - fewer voltage steps?
  - highest frequency at lowest V

Frequency/GHz	HS06	Power/W	HS06/GHz	HS06/W	Ratio to high
1.5	1085	286	723	3.79	98%
2.15	1424	330	662	4.32	111%
2.85	2032	524	713	3.88	100%

2) AMD node HEPSpec vs f
T.Hartmann, DESY

Carbon intensity of electricity grids
- Datacenter <u>location</u> matters



What about Storage?

- Around 40% of T2 energy used for storage
- ~10GB/s read/write, local+remote
  - 3PB RAID capable of 3 times this
- If 2PB is 90% spun down
  - factor 10 less energy with latency similar to tape
  - no robot, no winding, variable #'drives'
- Complete datasets on single disk
  - schedule BringOnline just like for tape
- Needs careful planning
  - disk failure probably whole dataset gone
    - can reproduce data
- Big benefits in cost and energy

Standard T2 disk

RAID 6, 12*10TB disk(€200), 100TB usable.
Server €10000 Euro
Bandwidth: 10 * 1GB/s
1PB 10*(10000+12*200) = 124k€/PB
Power: 10*(12*10+200)*8000/1000 =
25,600kWh/a

JBOD with spin-down

Server €5000. 100*10TB disk(€200), 10 active
Bandwidth: 1GB/s
1PB: 5000+100*200 = 25k€/PB
Power: (10*10+200)*8000/1000 = 2400kWh/a

TAPE

€8/TB, €5000/drive, Server(€10000)
Bandwidth: 1GB/s with 3 drives @ 300MB/s
1PB: 8000 = 8k€/PB, some drives and servers effectively dedicated. Robot.
Power: 40W/drive (200+3*40)*8000 =
2560kWh/a