

The challenge of powering the AI revolution



VENUE & DATE

**ICES Foundation Biennial
Workshop VIII**
October 1st, 2026

PRESENTED BY

Mario Paolone
Professor

OUTLINE

AI-driven data centers demand growth

Power grid constraints

Electricity markets integration

Local supply and GHG emissions

Opportunities: the EPFL HeatingBits project

Take home messages



OUTLINE

AI-driven data centers demand growth

Power grid constraints

Electricity markets integration

Local supply and GHG emissions

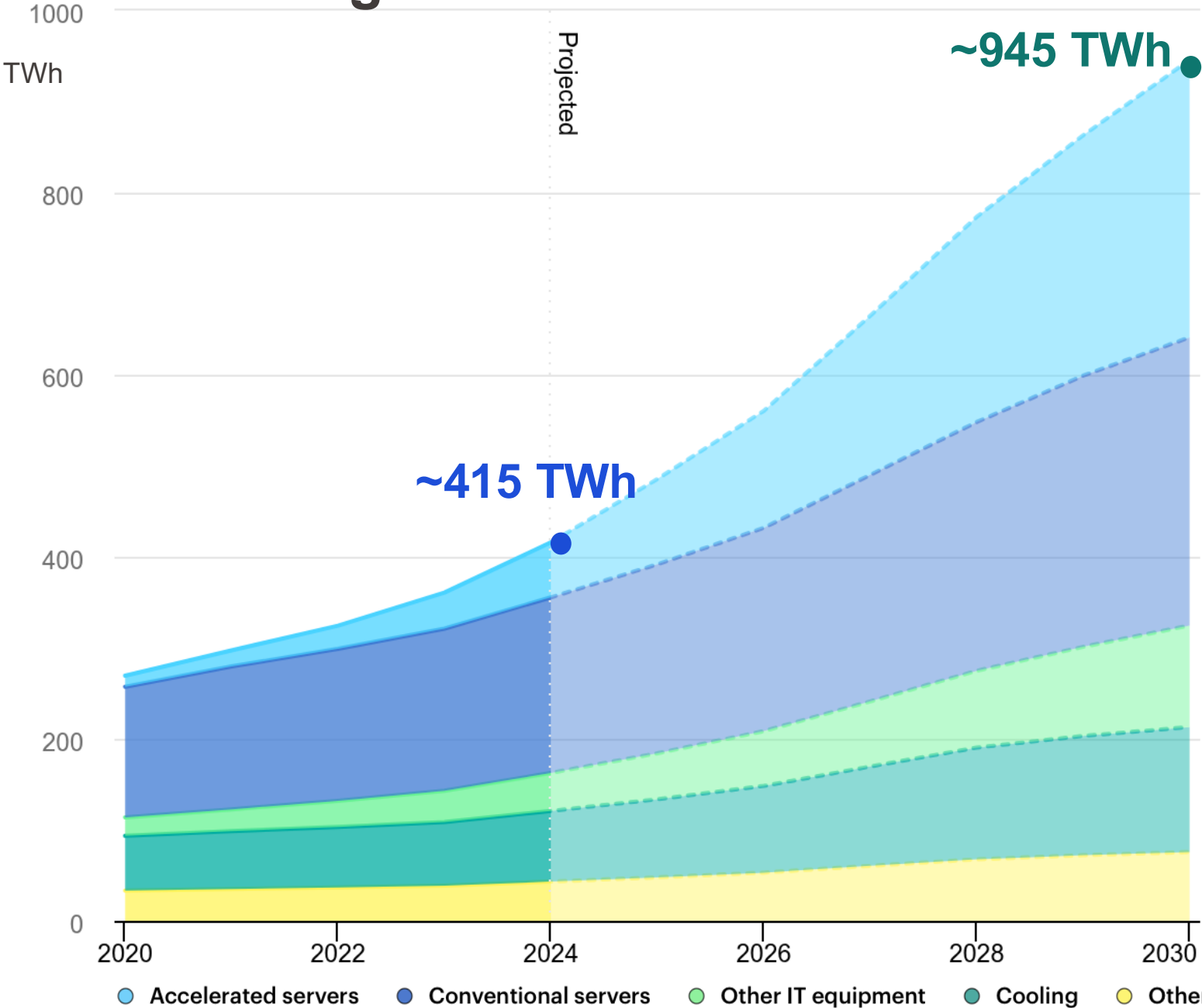
Opportunities: the EPFL HeatingBits project

Take home messages



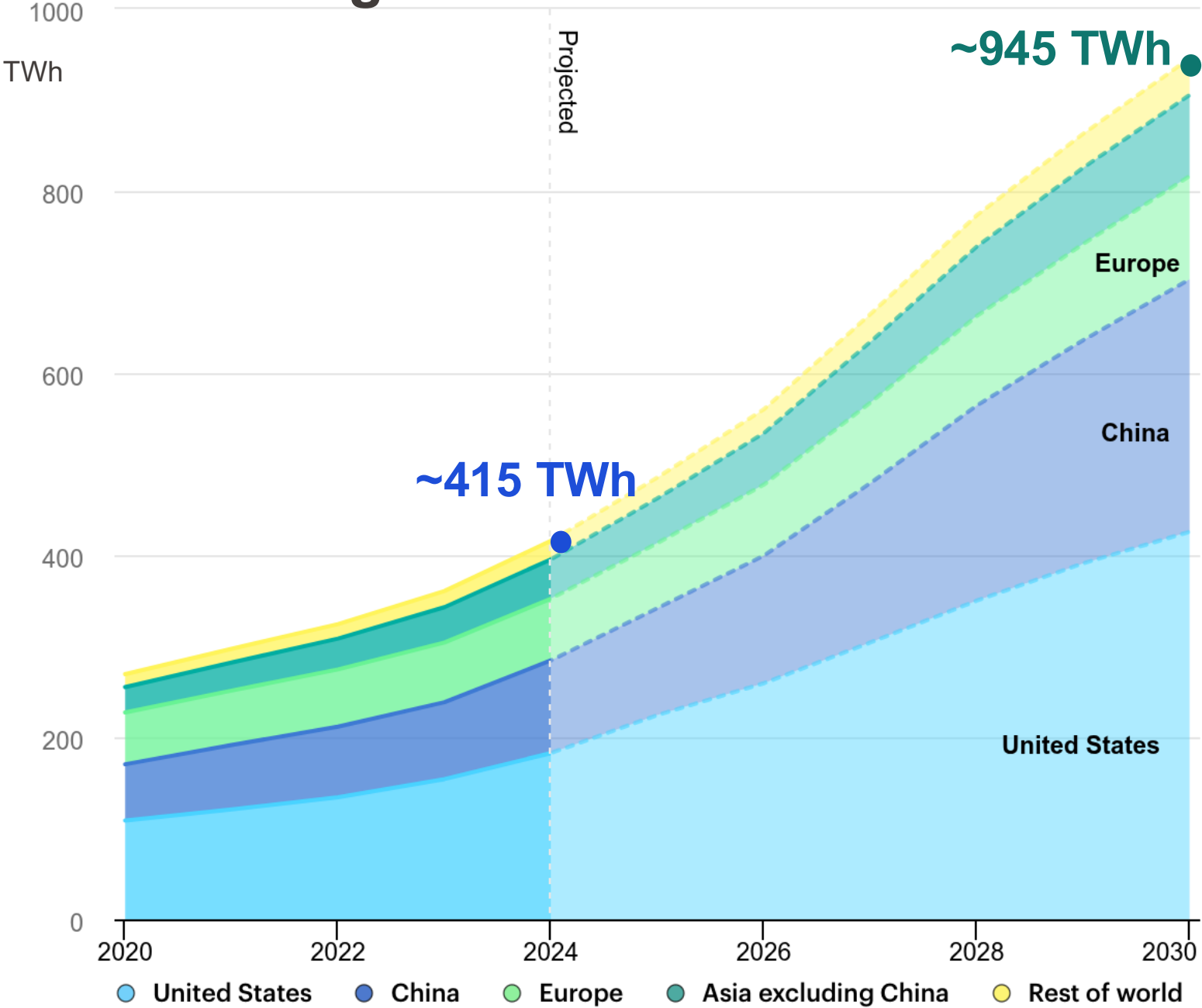
AI-driven data centers demand growth

IEA base case growth of global data centre electricity consumption by equipment.



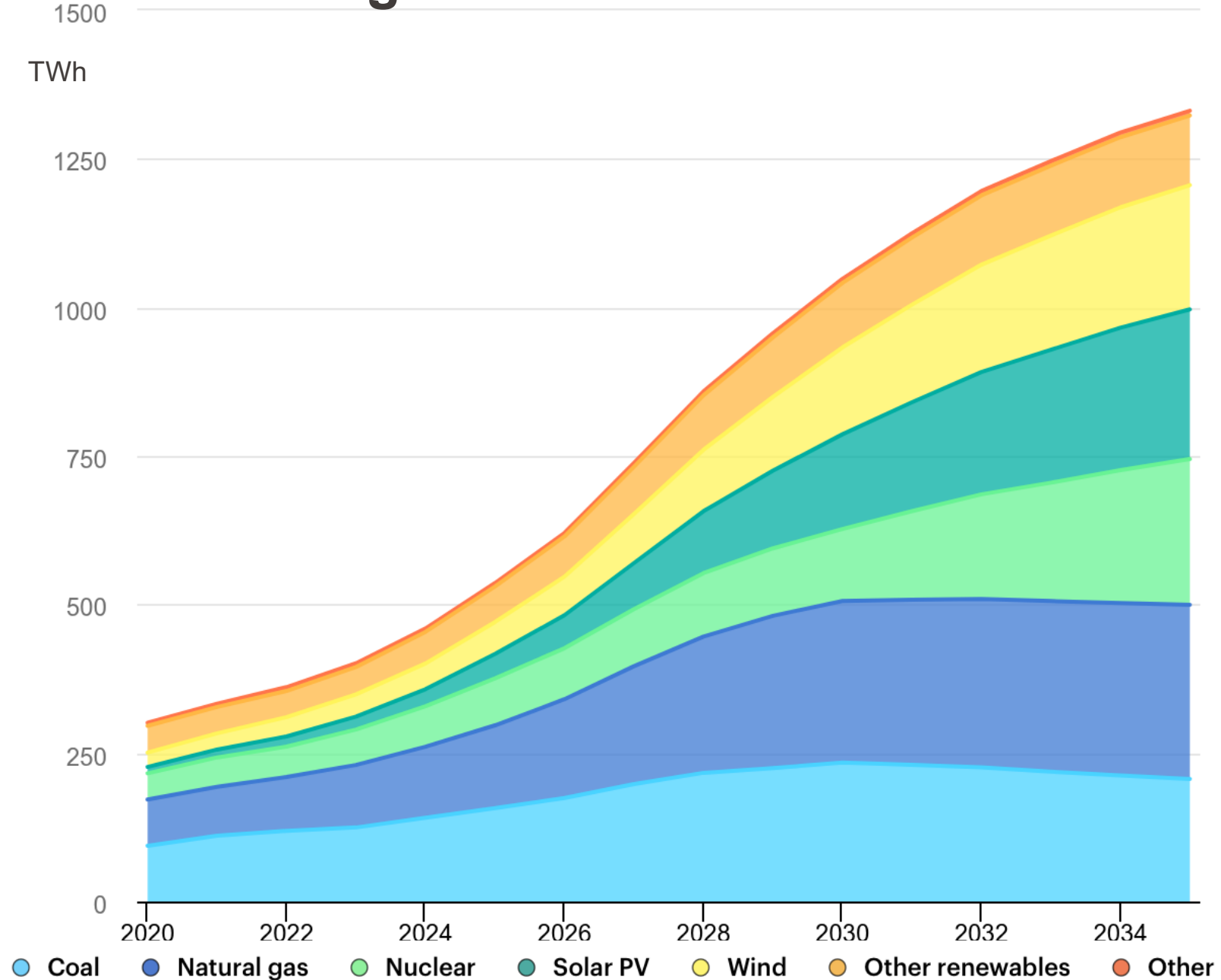
Source: IEA (2025), Global data centre electricity consumption, by equipment, Base Case, 2020-2030, IEA, Paris <https://www.iea.org/data-and-statistics/charts/global-data-centre-electricity-consumption-by-equipment-base-case-2020-2030>, Licence: CC BY 4.0

IEA base case data centers electricity consumption by region



Source: IEA (2025), Data centre electricity consumption by region, Base Case, 2020-2030, IEA, Paris <https://www.iea.org/data-and-statistics/charts/data-centre-electricity-consumption-by-region-base-case-2020-2030>, Licence: CC BY 4.0

IEA base case **sources of global electricity generation for data centers**



Source: IEA (2025), Sources of global electricity generation for data centres, Base Case, 2020-2035, IEA, Paris <https://www.iea.org/data-and-statistics/charts/sources-of-global-electricity-generation-for-data-centres-base-case-2020-2035>, Licence: CC BY 4.0

OUTLINE

AI-driven data centers demand growth

Power grid constraints

Electricity markets integration

Local supply and GHG emissions

Opportunities: the EPFL HeatingBits project

Take home messages



Why AI changes the load problem for the bulk power grid ?

- **Rapid investments** in large scale (GW-class) data centers

Classical cloud load

- Constant (known) deployment rate
- Many distributed workloads
- Moderate rack density (CPU-based)
- Demand associated to working hours patterns

AI data center

- Fast deployment time (potentially less than a year)
- GPU-based clusters with high utilization
- Demand driven by LLM training/querying patterns

- In modern data centers **power density** matters as much as energy needs

11×

AI-server **power density** increase
2020 → 2025

4×

Further increase projected
2025 → 2027

For the power grid, an AI data center is a geographically concentrated, high-load-factor campus that is similar to a large industrial plant, but with a much faster investment cycle.

EPFL Power grid constraints

Annual energy accounting

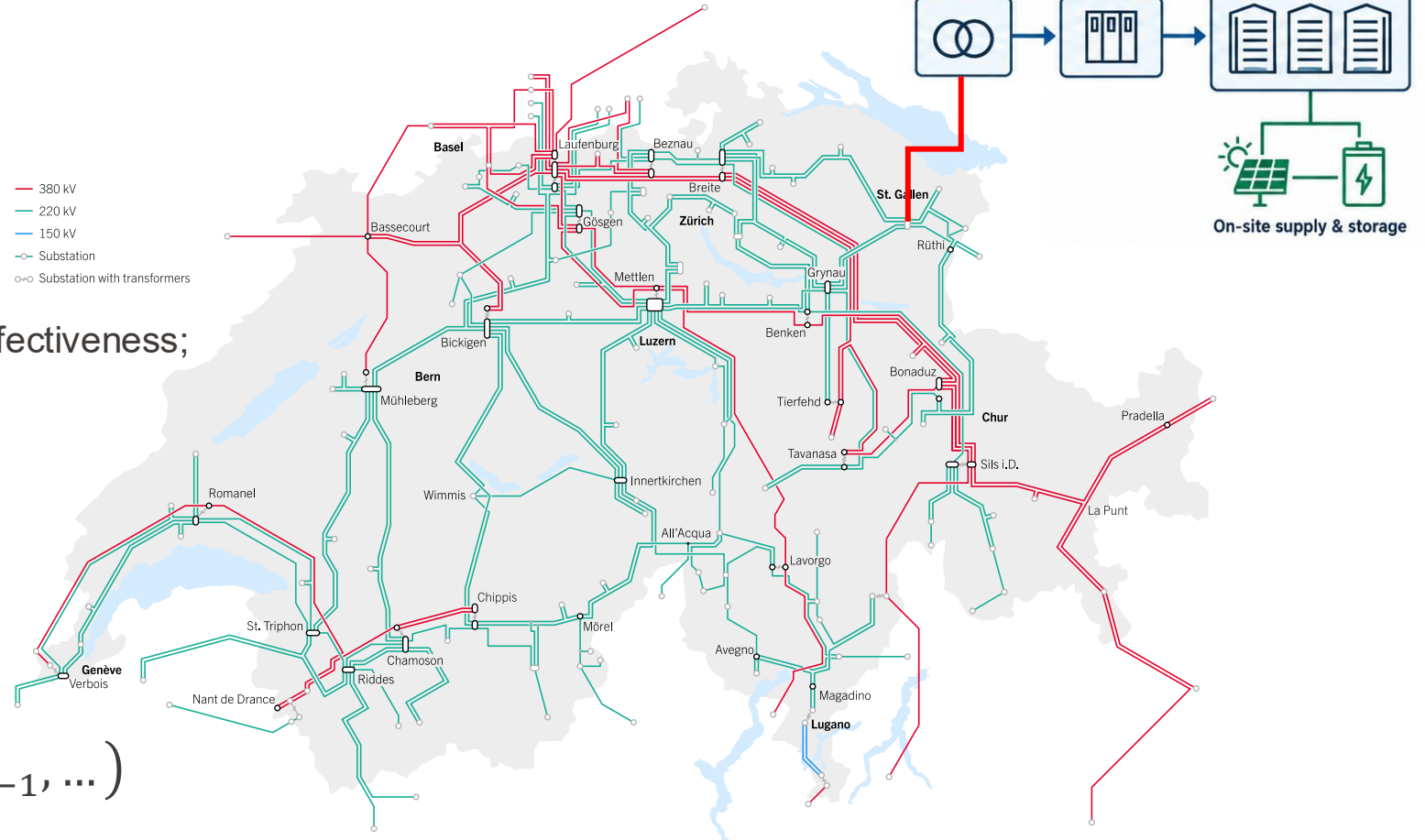
$$E_{DtC} = P_{DtC} \cdot PUE \cdot LF \cdot T$$

- P_{DtC} : data center power;
- $PUE = \frac{\text{Total DtC energy}}{\text{IT equipment energy}} \Big|_T$ power usage effectiveness;
- $LF = \frac{P_{DtC}^{avg}}{P_{DtC}^{peak}} \Big|_T$ load factor;
- T : time period (e.g. 1 year).

Grid hosting capacity limits

$$P_{node}^{max} \leq \min(P_{th}, P_V, P_{VS}, P_{fS}, P_{N-1}, \dots)$$

- P_{th} : thermal limit of the substation/line;
- P_V : power associated to min/max nodal voltage limits;
- P_{VS} : max power associated to grid voltage stability;
- P_{fS} : max power associated to frequency/angular stability;
- P_{N-1} : max power associated to N-1 contingency analysis



A country can have enough annual generation and still be unable to connect a GW-class data center at the desired node in the desired year.

Why AI changes the load problem for the bulk power grid ?

□ The infrastructure timescales are mismatched

AI servers / racks		Months	Commodity + accelerator supply chain dependent
Data-center shell		1–3 years	Civil construction can be modular and repeatable
Gas turbines / engines		2–5+ years	Equipment queues, permitting, fuel interconnection
Utility-scale renewables		1–5+ years	Interconnection queue and land/permitting dominate
Major transmission		5–15+ years	Planning, permitting, rights of way, supply chain

OUTLINE

AI-driven data centers demand growth

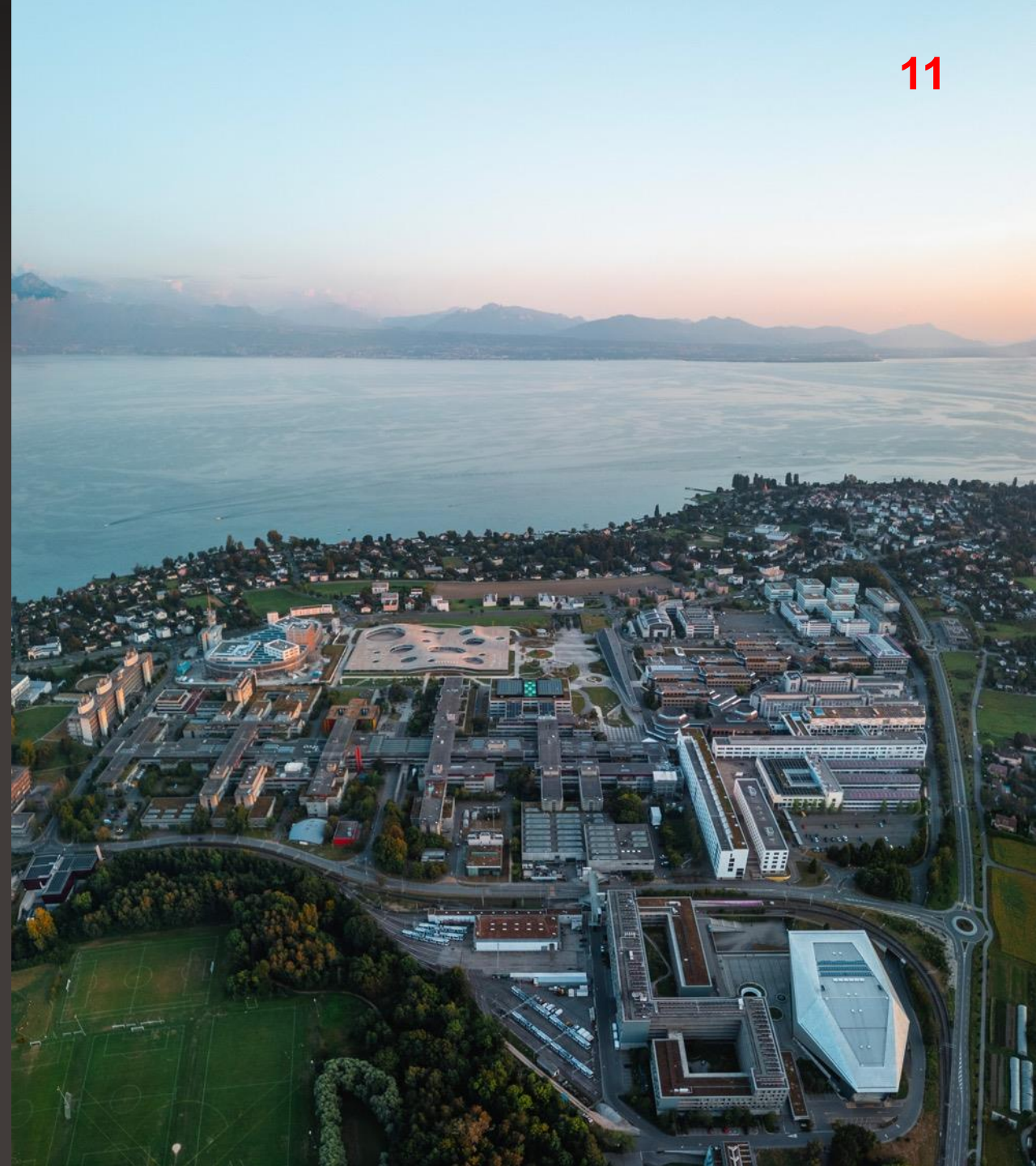
Power grid constraints

Electricity markets integration

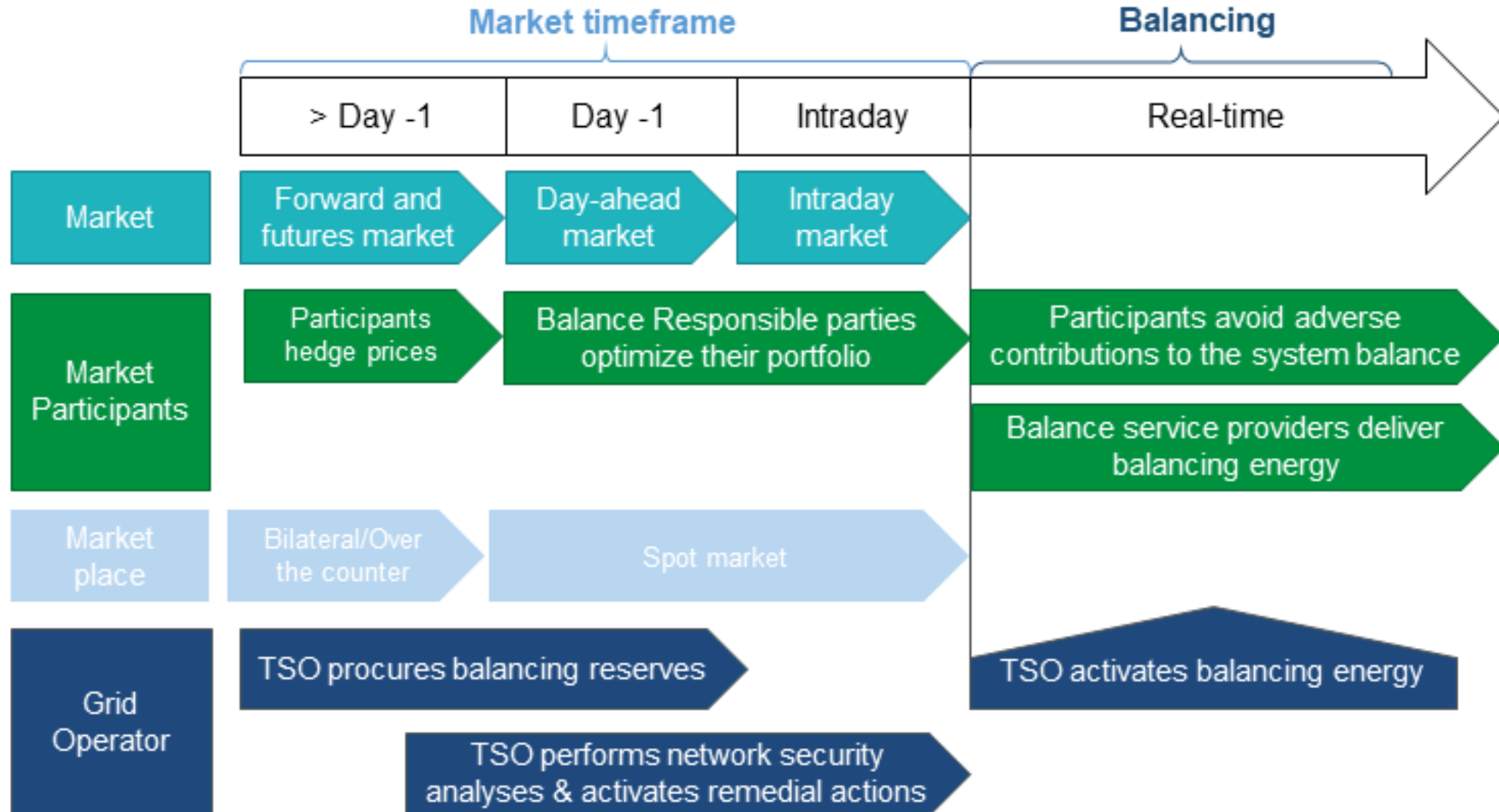
Local supply and GHG emissions

Opportunities: the EPFL HeatingBits project

Take home messages



Electricity markets timeline



Data centers can be passive loads or active participants

Passive loads

- Fixed MW request
- High load factor
- Backup assets
- Grid reinforcement treated as a utility

Active market participant

- Flexible power via dedicated power purchase agreements (PPAs)
- Batteries and thermal inertia as controllable assets
- Workload control across time and location
- Participation to energy and ancillary services markets

Response time, data center leverage and types of market

Response time	Data center leverage	Types of market
Seconds	Co-located batteries	Frequency containment reserve (FCR)
Minutes	Co-located batteries, DtC cooling setpoints, thermal inertia, workload control	Frequency restoration reserve (FRR)
Hours		Day ahead, intraday, congestion management
Days	Geographic workload shifting	Day ahead and congestion management

The question of opportunity costs associated to curtailing the workload execution

$$\min \sum_t (C_t P_t + C_t^{WL-delay} D_t + C_{SLA} V_t + C_t^{CO_2} E_t^{CO_2} - R_t F_t)$$

s. t.

Data center power constraints

Workload hard completion deadlines

Battery SoC and power constraints

Data center cooling constraints

Market service requirements hard constraints

- ☐ $C_t P_t$: cost of used electricity;
- ☐ $C_t^{WL-delay} D_t$: cost of delayed workloads;
- ☐ $C_{SLA} V_t$: cost of grid service violation (i.e., wrong dispatch, lack of FCR/FRR response, etc.);
- ☐ $C_t^{CO_2} E_t^{CO_2}$: carbon emissions costs;
- ☐ $R_t F_t$: revenues due to participation to electricity (energy and ancillary services) markets.

Day ahead market participation (EPFL DtC, Swiss days ahead market)

- Data Center buys energy on the DAM solving a **stochastic optimization problem**

$$\max \beta \text{Profit} + (1 - \beta) \text{Risk}$$

decision variables

Flexible workload schedule

Energy storage arbitrage (e.g. battery)

Local PV generation

Organic Rankine Cycle for heat recovery (if any)

s. t.

Data center power constraints

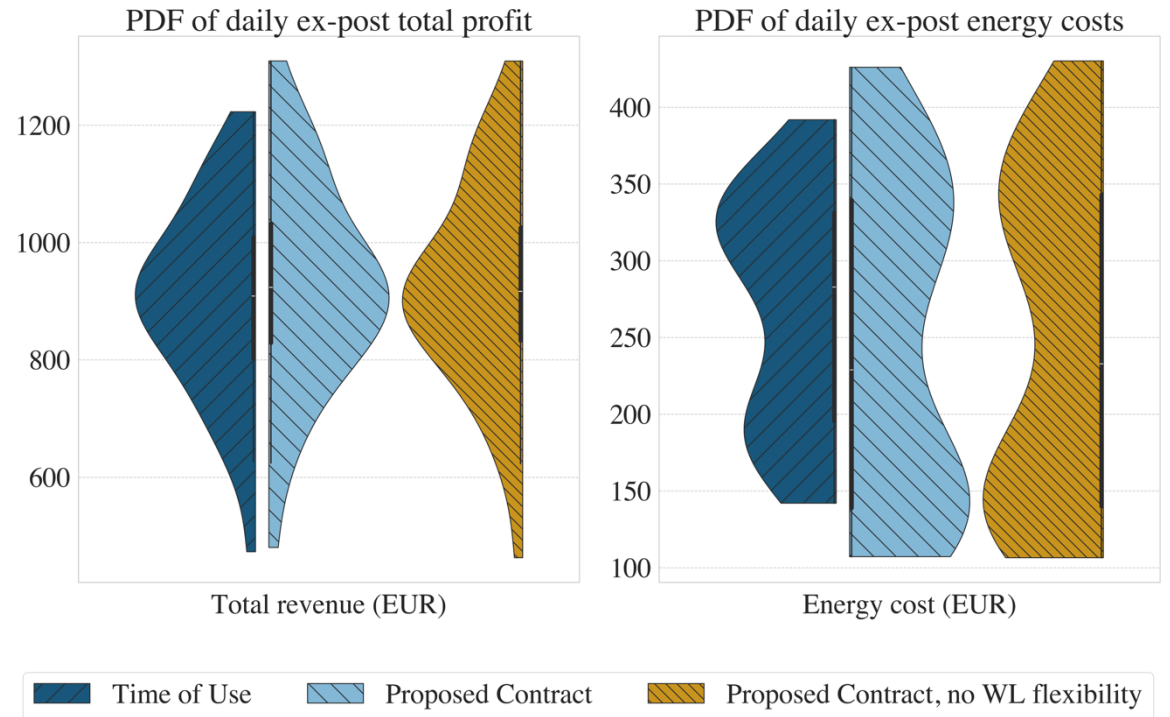
Workload hard completion deadlines

Battery SoC and power constraints

Data center cooling constraints

Market service requirements hard constraints

- Time of Use (ToU): typical ToU professional contract from Swiss DSO
- Use of real workload data from the Research Compute Center (RCP) of EPFL
- Electricity spot prices from Switzerland
- Ex post results from 40 days in 2025

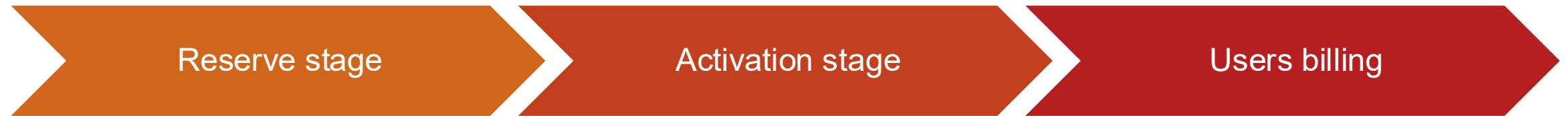


- Participation **does not significantly improve profits (+2.5%)**
- **Energy costs decrease by 10%**
- **Compute is more valuable than day ahead market opportunities**

Without impacting data center profits, their participation to day ahead markets is valuable only for grid operators as it improves data centers consumption predictability.

Balancing market participation

Use of powercap or CPU/GPU dynamic voltage and frequency scaling



□ Bid **availability**.

- Awarded bids determine
 - Share of flexible units, C_f
 - Baseline operation of flexible units, f_b

□ Baseline operation with (C_f, f_b)

- Bid **activation**
 - Upon activation signal, increase/decrease performance of flexible units f_{act}

□ Flexible users:

- Base price v_f with rebate v_{reb} upon activation

□ Rigid users:

- Base price v_r with capped surcharge v_{sc} upon activation

Market product	Data center action	Potential	Reason
Upward reserve	Availability to reduce consumption	High	Baseline operation is normal
Upward activation	Activate consumption decrease	Medium	Reduces compute output
Downward reserve	Availability to increase consumption	Low	Baseline operation is capped
Downward activation	Activate consumption increase	High	Increases compute output

Reserve provision cost C_{res}

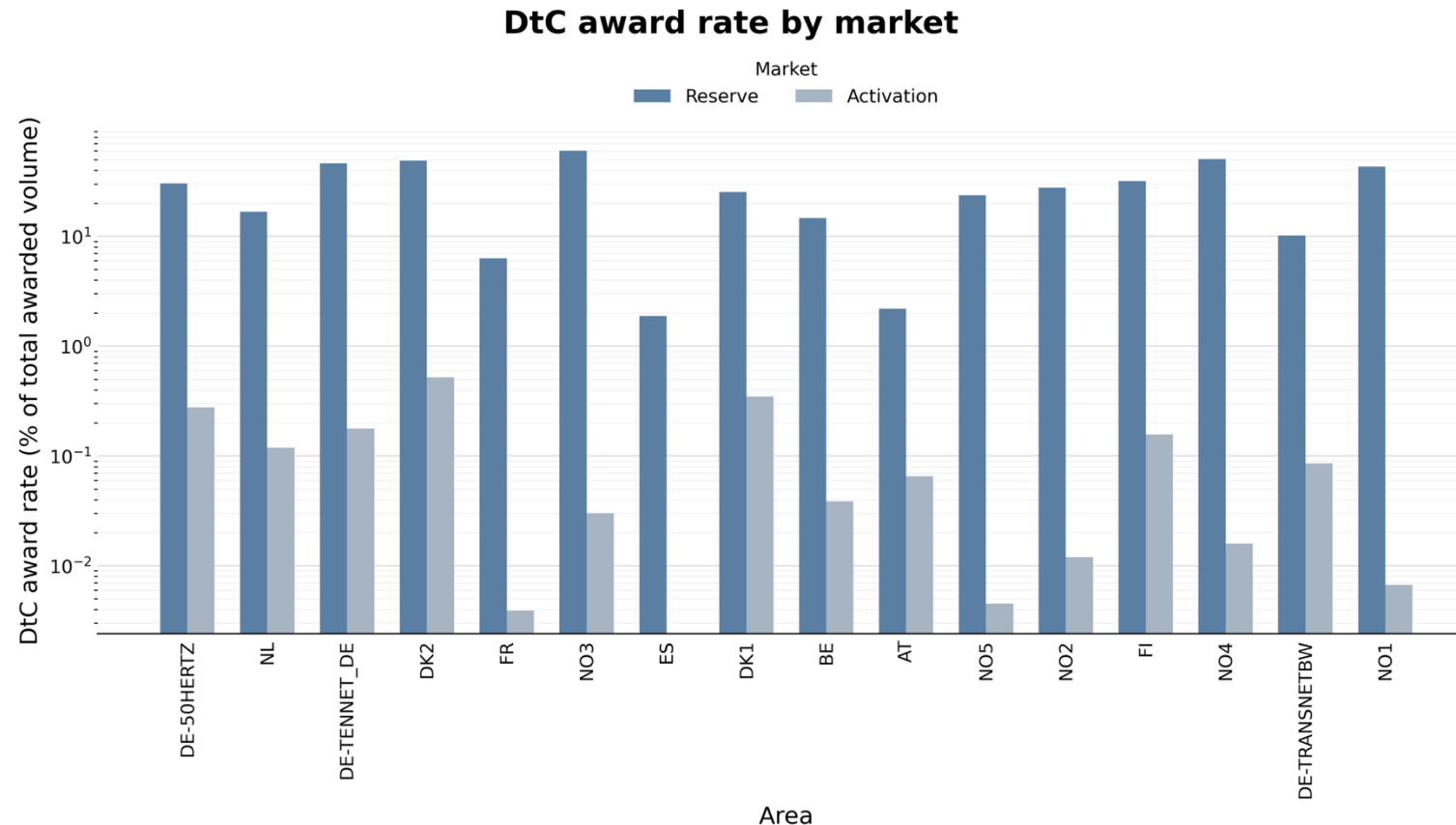
$$C_{res} \propto C_f(v_r - v_f)$$

Upward activation cost $C_{act}^{\uparrow}$

$$C_{act}^{\uparrow} \propto C_f \left(v_{reb}(1 - f_{act}) + v_f(f_{act}^{-1} - 1) \right) - (1 - C_f)(v_{sc}(1 - f_{act}))$$

Balancing market participation

- Simulated awarded bids in European markets with 2026 market data
- Upward balancing only
- Reserve participation > 10% of total volume in many regions (data centers are competitive)
- Activation participation < 1% of total volume across regions (data centers less competitive)



OUTLINE

AI-driven data centers demand growth

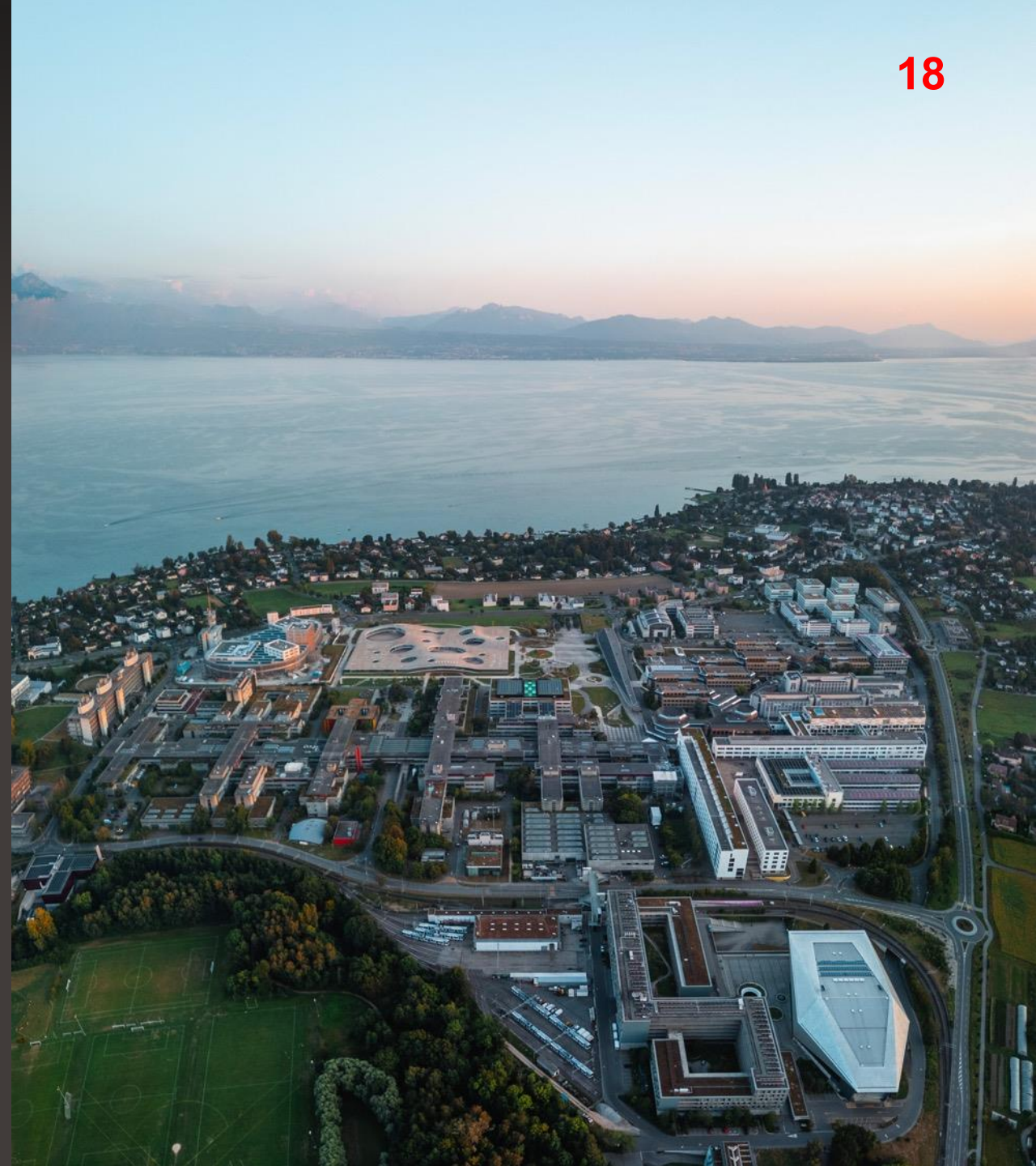
Power grid constraints

Electricity markets integration

Local supply and GHG emissions

Opportunities: the EPFL HeatingBits project

Take home messages



Power supply co-located with the data center

Option	Strength	Constraints
Grid + power purchase agreement	Scales through markets	Does not solve local hosting alone
Renewable co-location, battery	Low marginal emissions, BESS offers potential for day-ahead dispatch and ancillary services provision	Intermittency, land limits, grid export constraints
Gas fired generation	Firm and fast relative to the power grid	Fuel prices, GHG emissions

Local supply is not a single technology choice. It is a portfolio design problem under reliability, lead-time, emissions and interconnection constraints.

Why gas is attractive and why it is problematic

Attractive from the developer perspective

- Dispatchable power at high capacity factor
- Remove grid dependence/constraints
- Compatible with backup-power and reliability needs
- Commercial readiness: fuel + turbine + interconnection

Problematic from the system perspective

- Direct CO₂ emissions become coupled to AI demand
- Fuel supply chain issues/prices
- Can compete for gas infrastructure and capacity
- May lock in fossil assets if grid delays persist

1 GW data center entirely powered by gas-fired turbines requires 7-8 TWh/year of electricity with emissions in the range of 3-5 MtCO₂/year.

EPFL Local supply and GHG emissions

21

21

Current trends

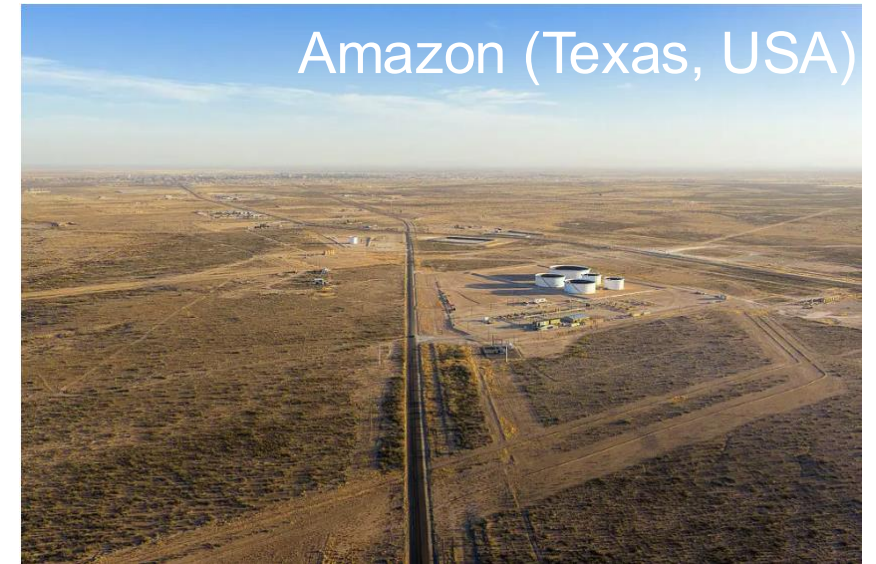


Fast deployment: 276 days from ground break to 200k GPUs

IT power: 946 MW (2GW at full expansion)

Power supply: **57 off-grid gas turbines**

Emissions: **Nitrogen oxide and Carbon monoxide emissions exceed by 48x and 81x the EPA norms**



Status: Under development

Power supply: **7.65GW** gas fired power plant

Emissions permitted: **33Mtons of CO₂eq annually**

OUTLINE

AI-driven data centers demand growth

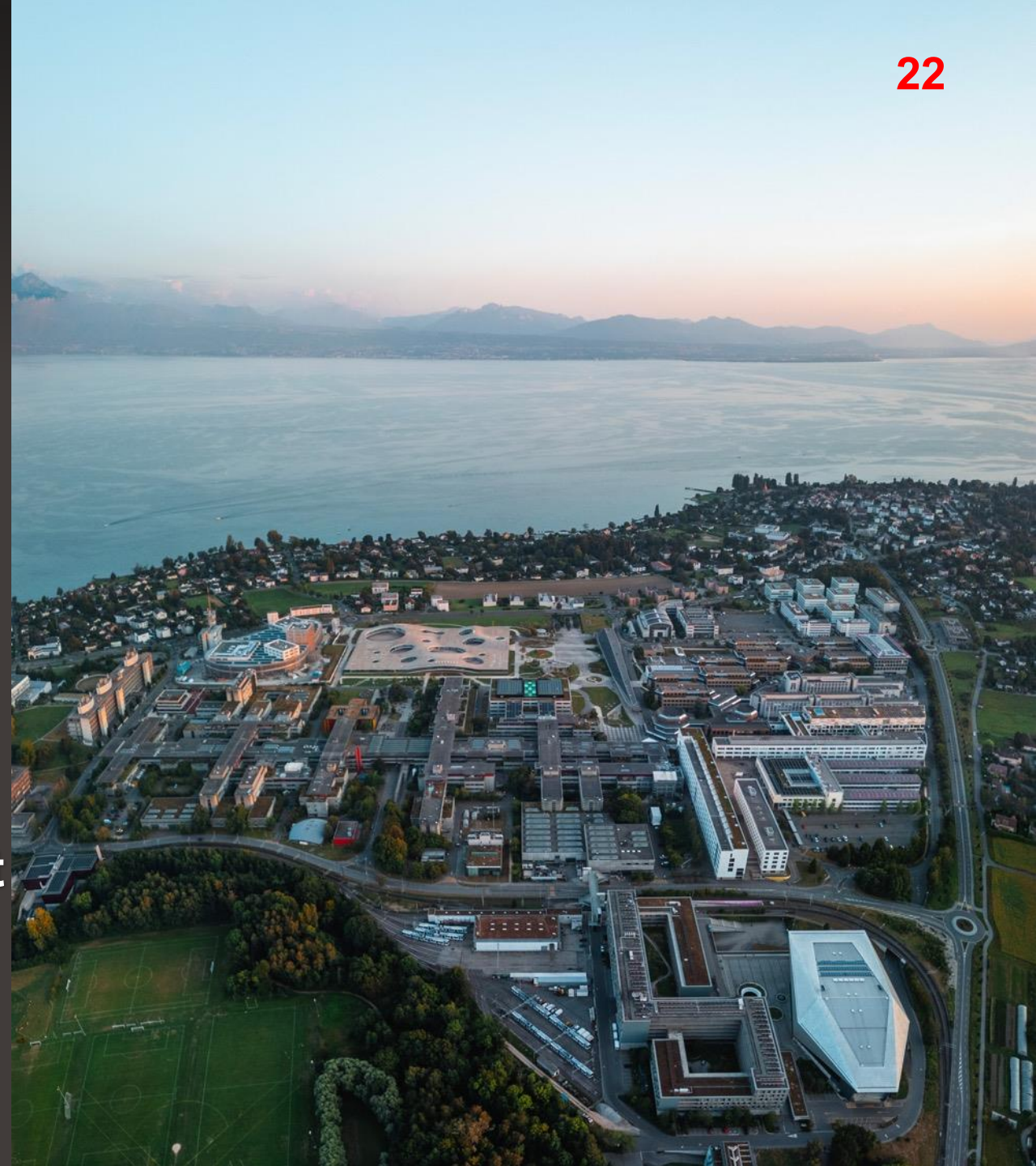
Power grid constraints

Electricity markets integration

Local supply and GHG emissions

Opportunities: the EPFL HeatingBits project

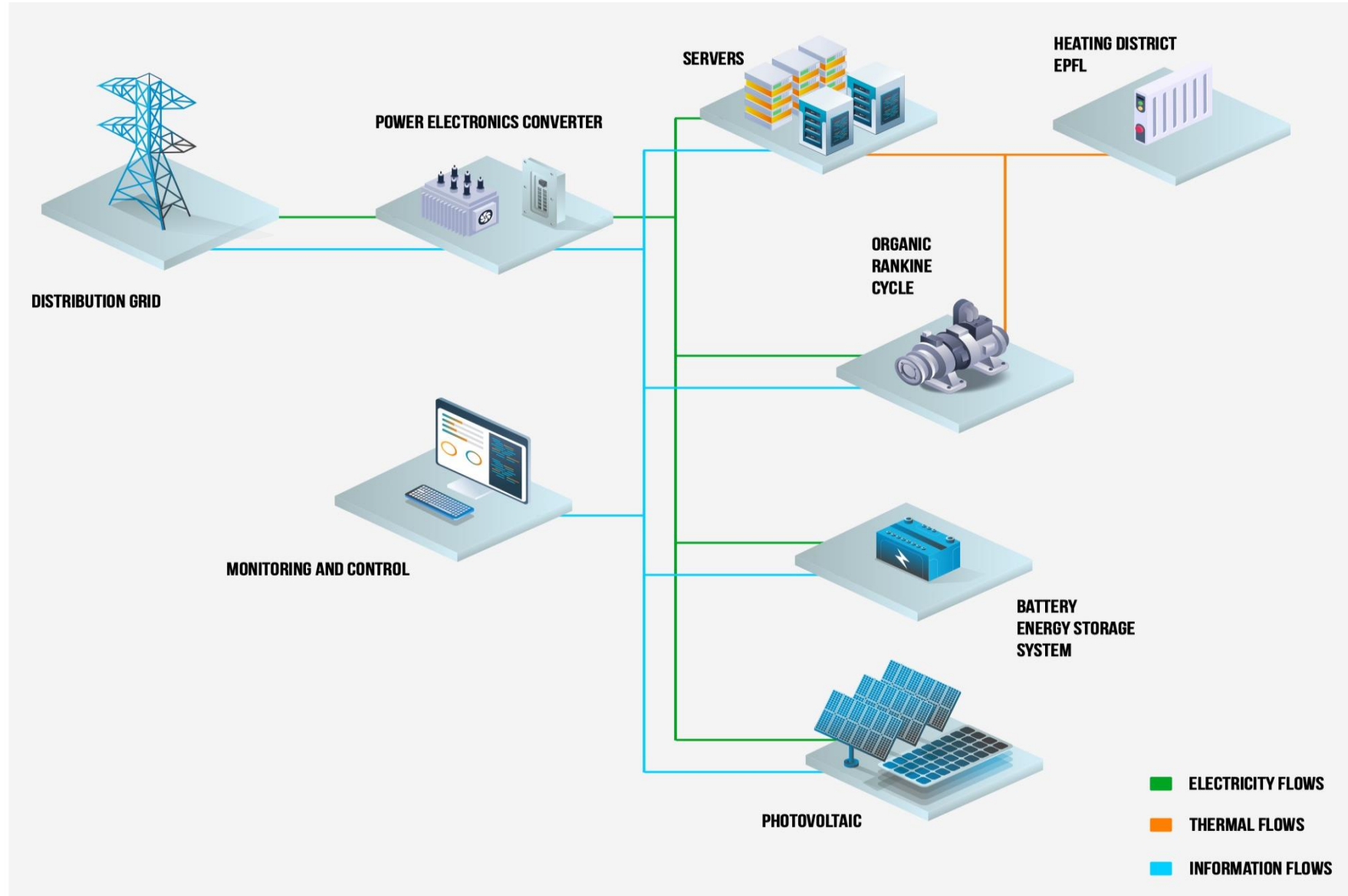
Take home messages



EPFL Opportunities: the EPFL HeatingBits project

23

Overview of the HeatingBits demonstrator



EPFL Opportunities: the EPFL HeatingBits project

24

Optimal cooling for heat extraction (courtesy of Prof. E. Matioli)

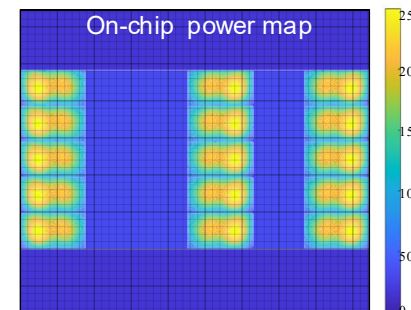
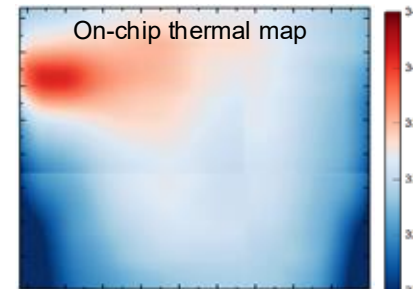
Outcome

- Accurate **heat map measurement** of the CPU, supported by experimental performance benchmarking
- Power-map-guided optimized **microfluidic design framework** for high-power CPUs or GPUs
- Systematic **integration of liquid-cooled CPU** in server
- **Demonstration** of high-temperature heat extraction

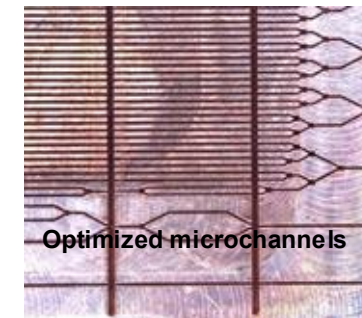
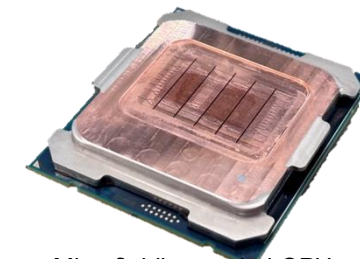
Impact

- Achieve heat **extraction efficiency** of 80%
- Achieve outlet coolant temperatures of approximately **75°C**
- Provides an **end-to-end cooling design framework** that can be applied to different high-power processors and support the reuse of decommissioned CPUs.

Thermal/Power map



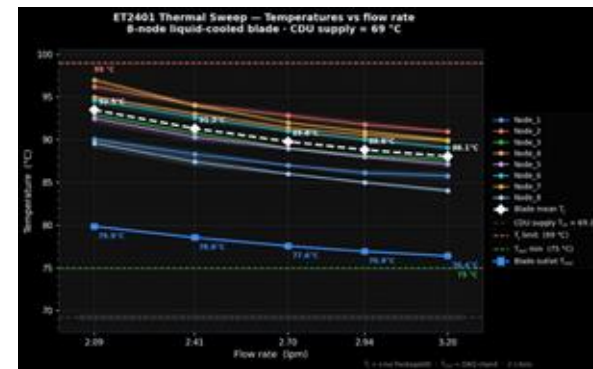
Cooling design



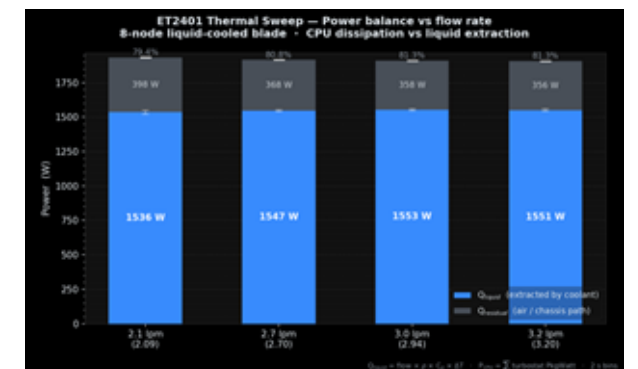
Server integration



Thermal performance



Heat extraction efficiency



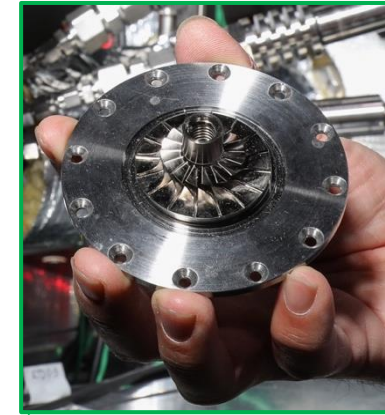
Waste heat recovery (courtesy of Prof. J. Schiffmann)

Outcome

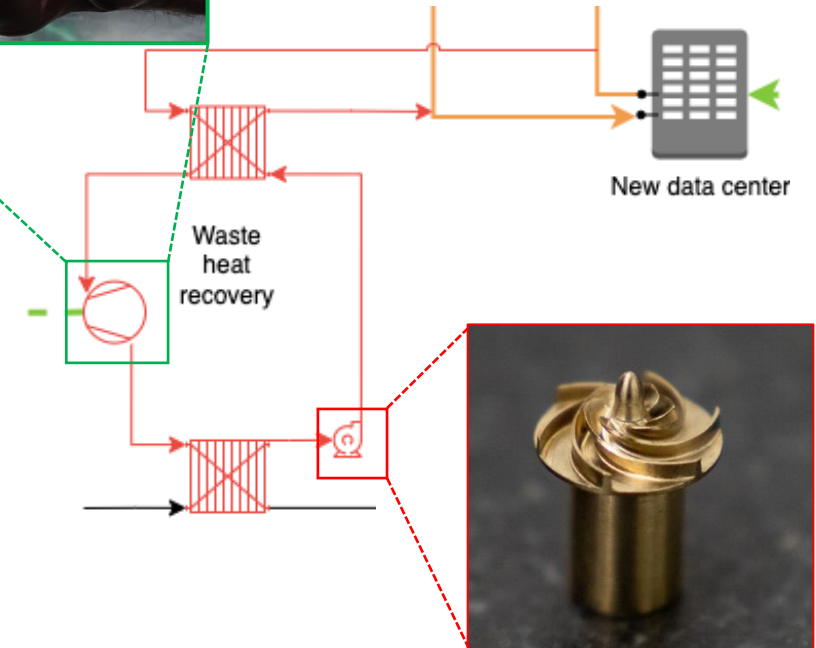
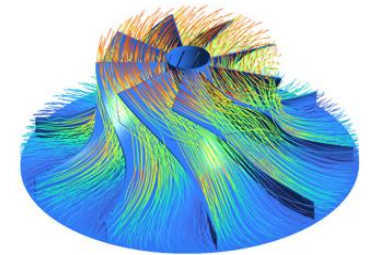
- Optimized **organic Rankine cycle** using R1233zd(E), selected based on performance, safety, and sustainability criteria.
- Adaptive variable-geometry **turbine** using data-driven models that outperform conventional mean-line predictions.

Impact

- Delivery of the first compact, fully **oil-free small-scale ORC**: ~4.7 kW net output, ~9% thermal efficiency and ~60% exergy efficiency.
- At EPFL data-center scale, **80% of recovered heat** supplies the heating network and **~10% is converted back into electricity**.
- Variable turbine geometry and data-driven design improve recovery under fluctuating loads while accelerating and **improving design accuracy**.



CFD simulation of ORC turbine



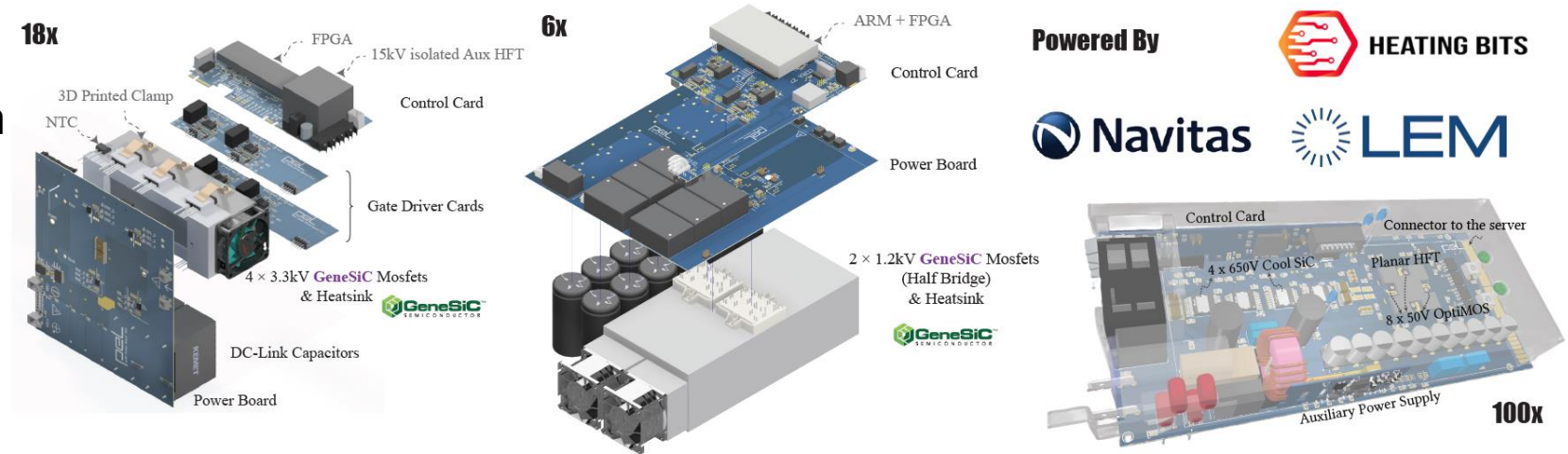
EPFL Opportunities: the EPFL HeatingBits project

26

Novel conversion solution for 800V data centers (courtesy of Prof. D. Dujic)

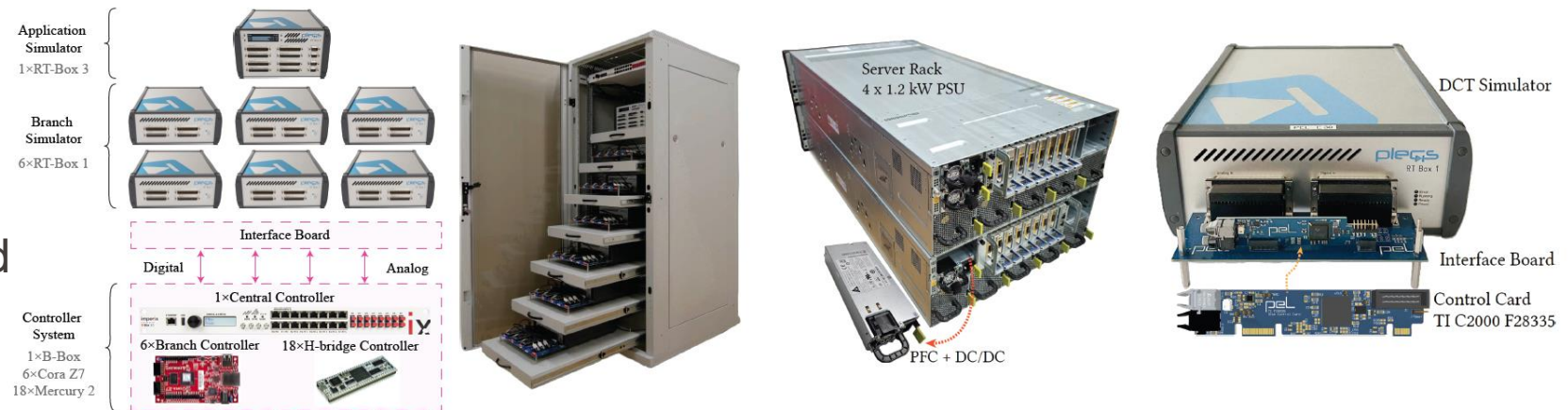
Outcome

- HVDC AC/DC supply based on a **solid-state transformer**, with optimized power hardware, magnetics and solid insulation.
- **Compact 800 V, 12 V converters** per server blade.



Impact

- Reduce conversion stages and improves **power-delivery efficiency to approximately 98%**.
- Facilitate the **integration of batteries and DC sources**, including photovoltaic generation.
- Increase overall **power-system reliability**.

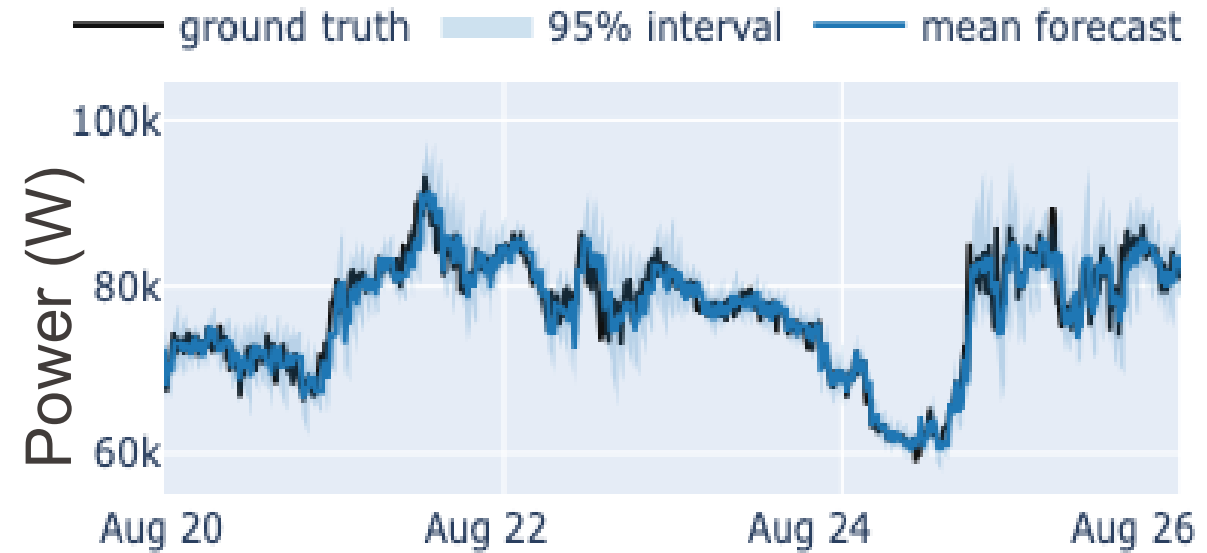


Workload Characterization & Power Forecasting (courtesy of Prof. D. Atienza)

Outcome

- **Black-box performance models**, CloudProphet, GreenDVFS, and CloudFormer, for predicting application behavior and optimizing processor operation without access to internal workload data.
- Curated more than **one year of GPU power data** and developed an hour-ahead foundation model for GPU power forecasting.

Hour-Ahead GPU Power Consumption Forecast



Impact

- CloudProphet **reduced** black-box **performance-prediction errors** by approximately **50%**.
- GreenDVFS **reduced energy consumption by 18%** and **processor temperatures by 30%**, while CloudFormer achieved a **7.8% mean absolute error on unseen workloads**.
- The hour-ahead forecasting model achieved a **CRPS of 0.14**.

Optimal planning of data center co-located PV and batteries

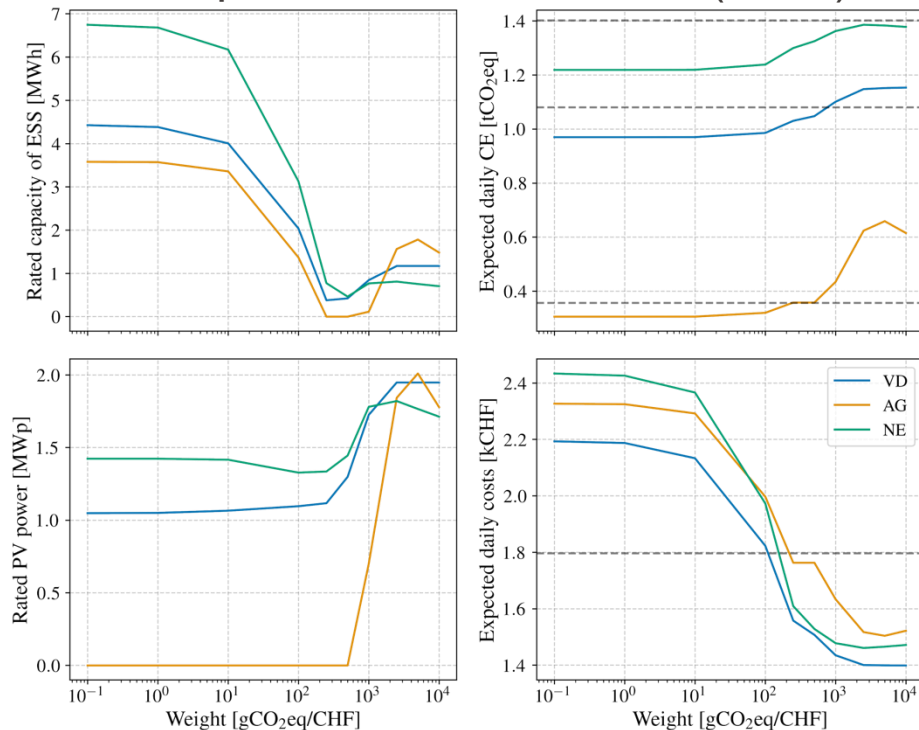
Long-term planning for sustainable and grid-friendly DCs

- Ensure dispatchability
- Sizing of battery and local electricity generation from PV
- Minimize **carbon footprint** and **OPEX**

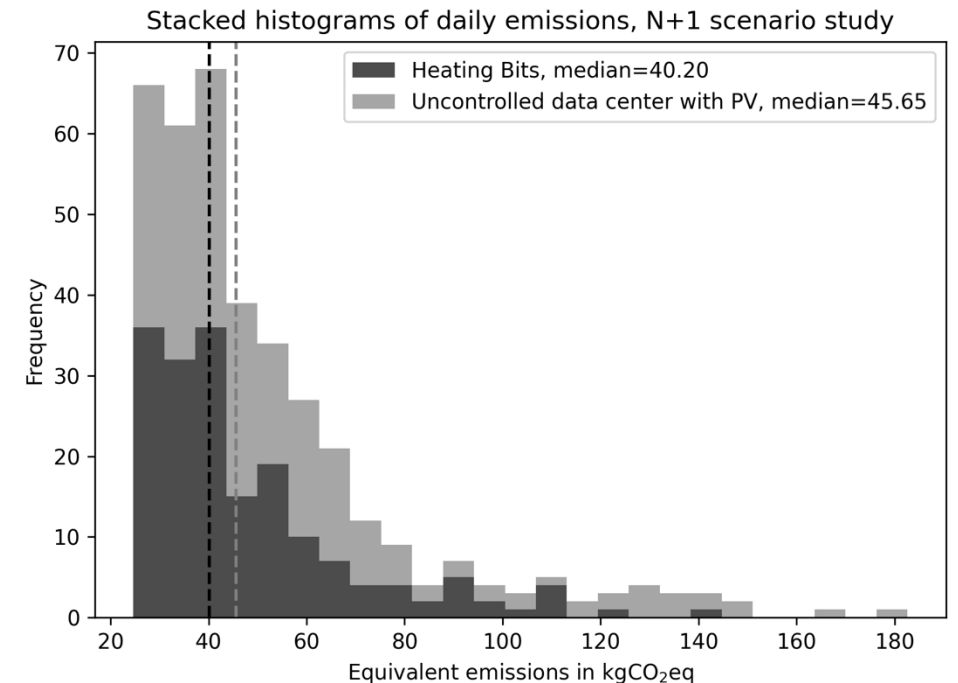
Example of results

Long-term planning (for Switzerland)

- Max. 15% expected reduction of emissions (Neuchâtel)
- Max. 22% expected costs reduction (Vaud)



- Substantial CO₂ footprint reduction
- System stochasticity reduced by 60%

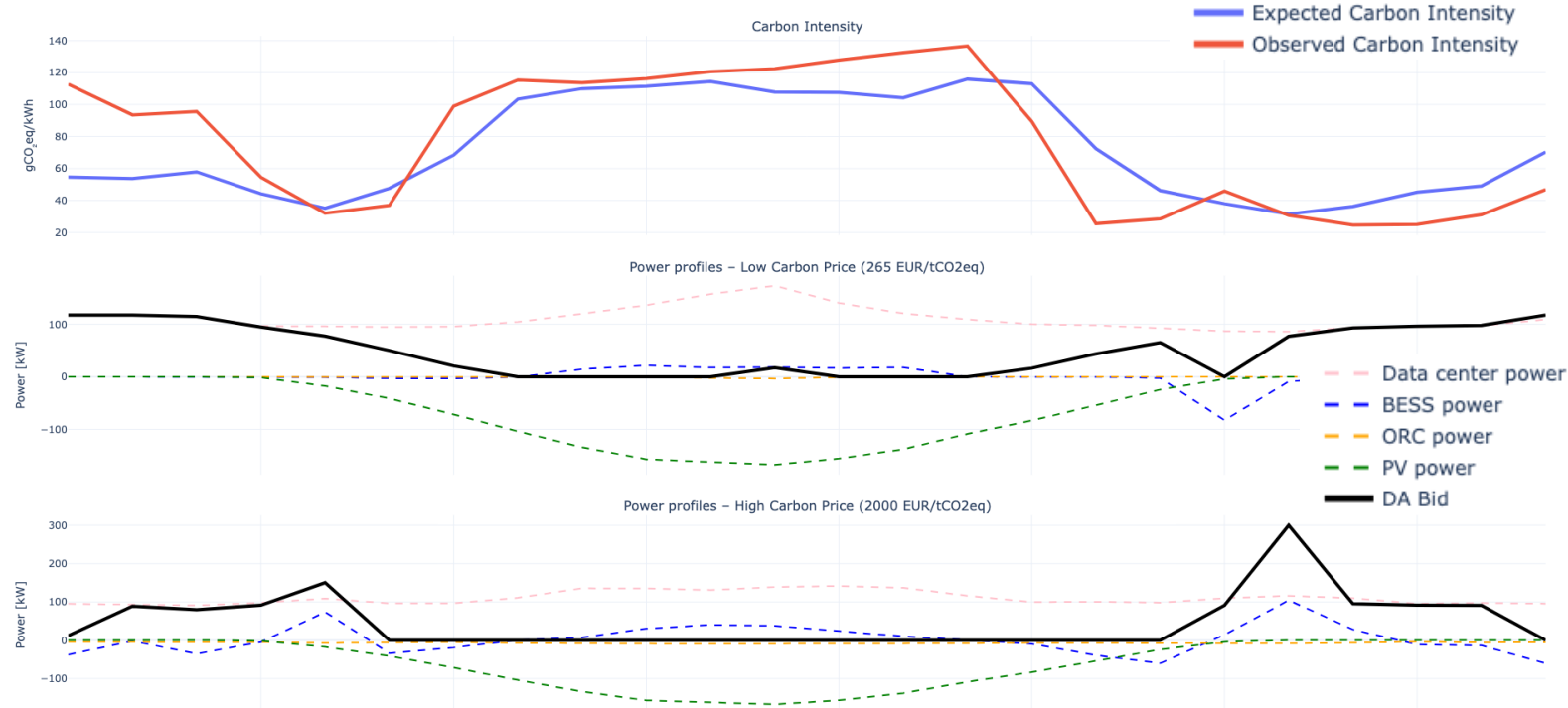


EPFL Opportunities: the EPFL HeatingBits project

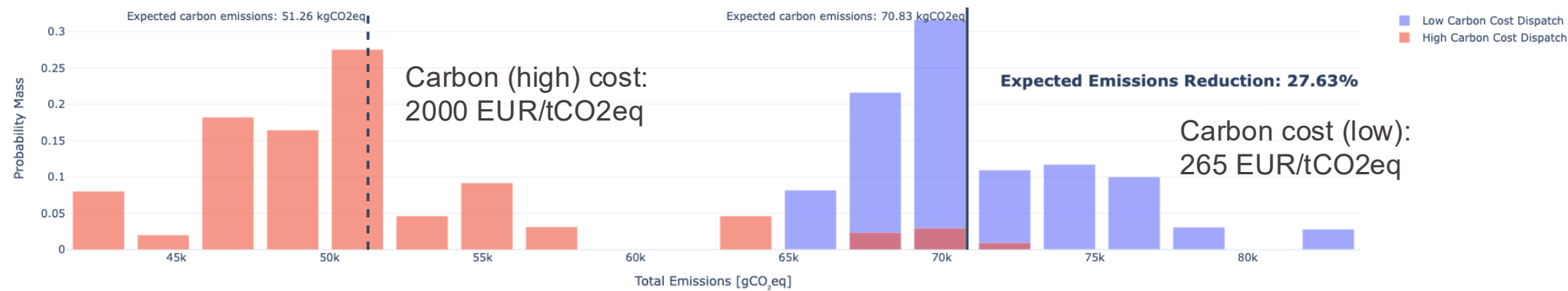
Optimal operation example

29

Day-Ahead Dispatches: Expected Operation with Different Strategies 2025-07-01



Probability Distribution Functions of Carbon Emissions 2025-07-01



Example of operation of the EPFL RCP datacenter: DC 200kW
PV 200kW, BESS 250kW/kWh
ORC 10kW.

Day-ahead bidding strategy

Bid on spot market, leveraging flexible assets accounting for:

- spot electricity prices;
- Grid imbalance costs;
- operational carbon emissions through carbon cost;
- risk aversion.

OUTLINE

AI-driven data centers demand growth

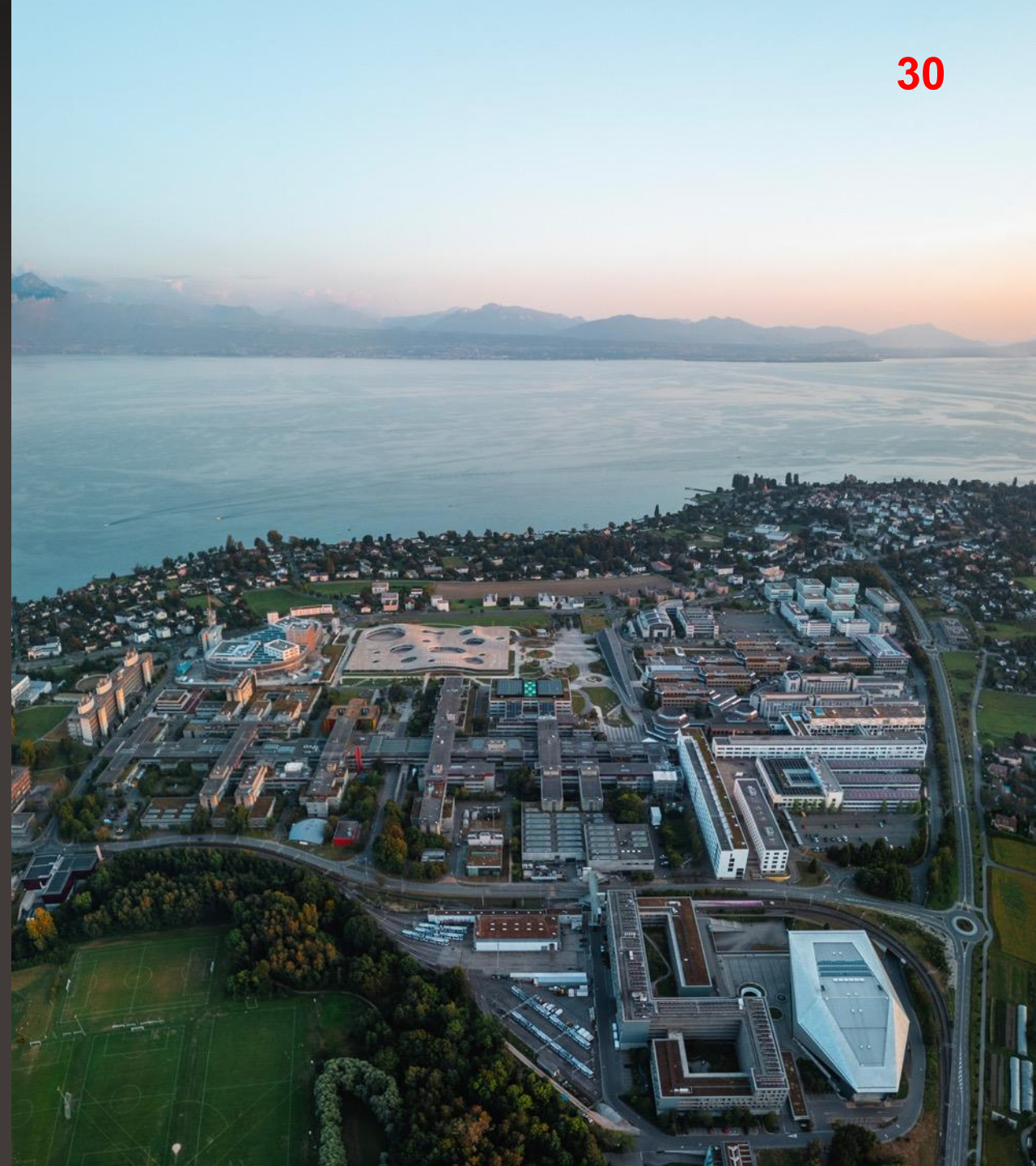
Power grid constraints

Electricity markets integration

Local supply and GHG emissions

Opportunities: the EPFL HeatingBits project

Take home messages



- Electricity demand of AI-driven data centers is becoming substantial, and its impact has **system-level consequences**.
- The immediate bottleneck is often **grid hosting capacity**: transformers, substations, transmission stability and N-1 security.
- Data centers can be active electricity market participants, but compute flexibility must be quantified as current evidences show very small savings vs computation revenues. In this respect, specific **grid codes** (under development on many countries) **may impose specific rules for their integration into the power grid**.
- Local supply can accelerate deployment; **gas solves firmness but creates significant GHG emissions**.
- **The durable solution is coordinated planning across compute, grid, markets and low-carbon supply.**

Powering AI is no longer only an energy-procurement problem. It is a system optimal planning and operation problem.

The EPFL logo is displayed in red, bold, sans-serif capital letters in the top left corner of the slide.

EPFL

The background of the slide is a photograph of the EPFL campus at dusk. A large, three-dimensional red sculpture of the letters 'EPFL' sits on a long, low, metallic rectangular base in a grassy area. In the background, there are trees and a modern building with a curved glass facade under a twilight sky.

Thank you for your attention !

■ École
polytechnique
fédérale
de Lausanne

Distributed
Electrical Systems
Laboratory

PRESENTED BY

Mario Paolone
Professor