EUROHPC SUMMIT 2024

ANTWERP

How Does One Define an "Energy Efficient" HPC System? From Data to Action

Andreas Wierse, Julita Corbalan,
Jean-Olivier Gerphagnon, Bastian Koller,
Per Öster, Ondřej Vysocký

UNLEASHING THE POWER OF EUROPEAN HPC AND QUANTUM COMPUTING

ANNOTATION

Our Panelists

- Andreas Wierse, *SICOS BW GmbH (Moderator)*
- Ondřej Vysocký, *IT4Innovations*
- Jean-Olivier Gerphagnon, *Eviden*
- Per Öster, *CSC – IT-Center for Science*
- Julita Corbalan, *Barcelona Supercomputing Center (BSC)*
- Bastian Koller, *High Performance Computing Centre Stuttgart (HLRS)*

ETP 4 HPC
EUROPEAN TECHNOLOGY PLATFORM FOR HIGH PERFORMANCE COMPUTING

Our Topic

How Does One Define an "Energy Efficient" HPC System?

From Data to Action!

Representatives of the European HPC value chain - HW, SW providers, infrastructure providers and organisations or companies involved in measuring the environmental footprint of HPC systems come together in this session to answer the following questions:

The Questions

- What else apart from energy consumption should be measured, and to what detail?
- To what extent should one look at factors other than energy consumption (e.g. SW) and include the broader environmental impact, spanning from the extraction of raw materials for component manufacturing to the expenses associated with decommissioning?
- What is the minimum set of data or criteria that a reliable measurement should include?
- Apart from the data-level, what are the expectations from the SW stack in terms of SW-level monitoring?
- Are there any successful use cases of systems where such a model has been implemented?
- Are there any existing examples of metrics that could serve as the skeleton of a standard?
- Is there any simple way to establish a measurement that would enable comparisons?
- Once such a model is established, how detailed should it be?
- What approximate level of granularity should it be based on?

The Questions

- What else apart from energy consumption should be measured, and to what detail?
- To what extent should one look at factors other than energy consumption (e.g. SW) and include the broader environmental impact, spanning from the extraction of raw materials for component manufacturing to the expenses associated with decommissioning?
- What is the minimum set of data or criteria that a reliable measurement should include?
- Apart from the data-level, what are the expectations from the SW stack in terms of SW-level monitoring?
- Are there any successful use cases of systems where such a model has been implemented?
- Are there any existing examples of metrics that could serve as the skeleton of a standard?
- Is there any simple way to establish a measurement that would enable comparisons?
- Once such a model is established, how detailed should it be?
- What approximate level of granularity should it be based on?

Mentimeter

Warm-Up questions:

- Who is here?
- What efficiency metrics of the system you use/operate are tracked?
- In what granularity are the systems measured that you typically use/operate?

Join at menti.com | use code 3876 9437

EUROHPC
SUMMIT 2024

ANTWERP
18–21 MARCH

TO EXASCALE
AND BEYOND

Ondřej Vysocký

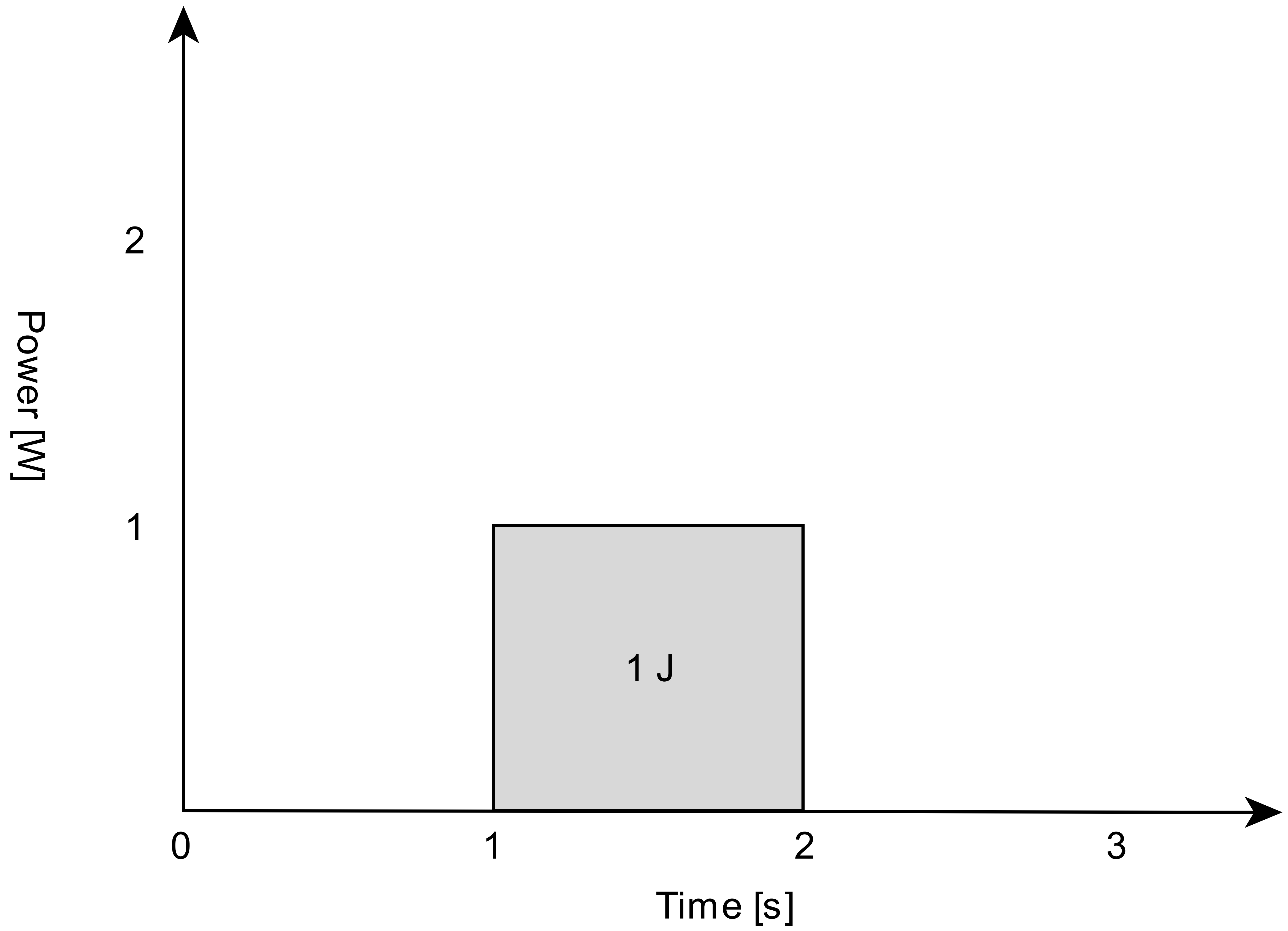
IT4Innovations

EuroHPC Joint Undertaking

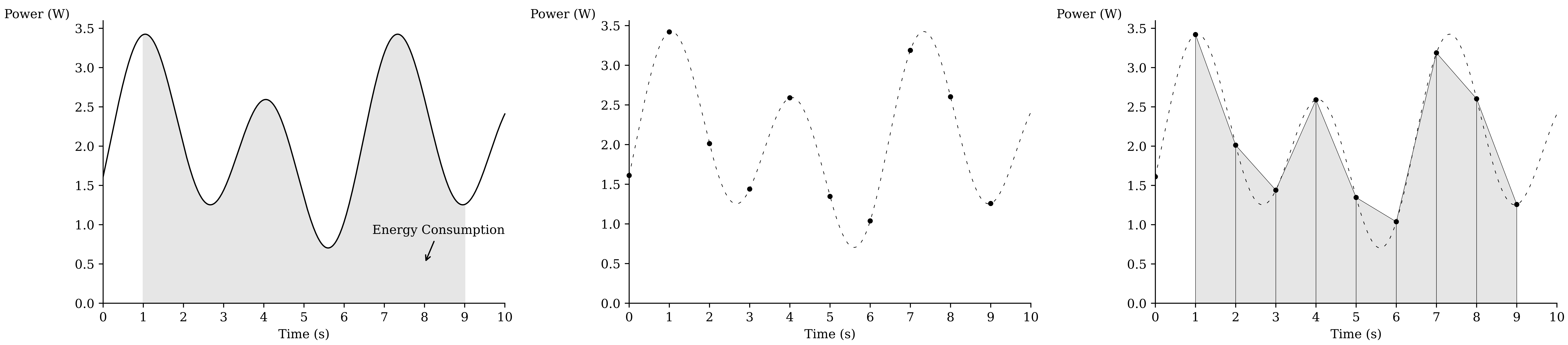
be EU belgium24.eu

$$Energy = Power \times Time$$

- Power [W]
- 1 W * 1 s = 1 J
- 1 W * 1 h = 1 Wh = 3 600 J



$$Energy(t) = \int_0^t Power(x)\,\mathrm{d}x \approx \frac{\sum_{i=0}^{n} PowerSample_i}{SamplingFrequency}$$



Img source, Luís Cruz (TU Delft)

Energy efficiency

THE GREEN 500™

- FLOPs/W

Green500 Data

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)
1	293	**Henri** - ThinkSystem SR670 V2, Intel Xeon Platinum 8362 32C 2.8GHz, NVIDIA H100 80GB PCIe, Infiniband HDR, Lenovo Flatiron Institute United States	8,288	2.88	44	65.396
2	44	**Frontier TDS** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	120,832	19.20	309	62.684

Energy efficiency

- FLOPs/W
- Rpeak/W

Green500 Data

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/ watts)
1	293	**Henri** - ThinkSystem SR670 V2, Intel Xeon Platinum 8362 32C 2.8GHz, NVIDIA H100 80GB PCIe, Infiniband HDR, Lenovo Flatiron Institute United States	8,288	2.88	44	65.396
2	44	**Frontier TDS** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	120,832	19.20	309	62.684

Energy efficiency

- FLOPs/W
- Rpeak/W
- LUPs/W
  (Lattice Boltzmann methods)

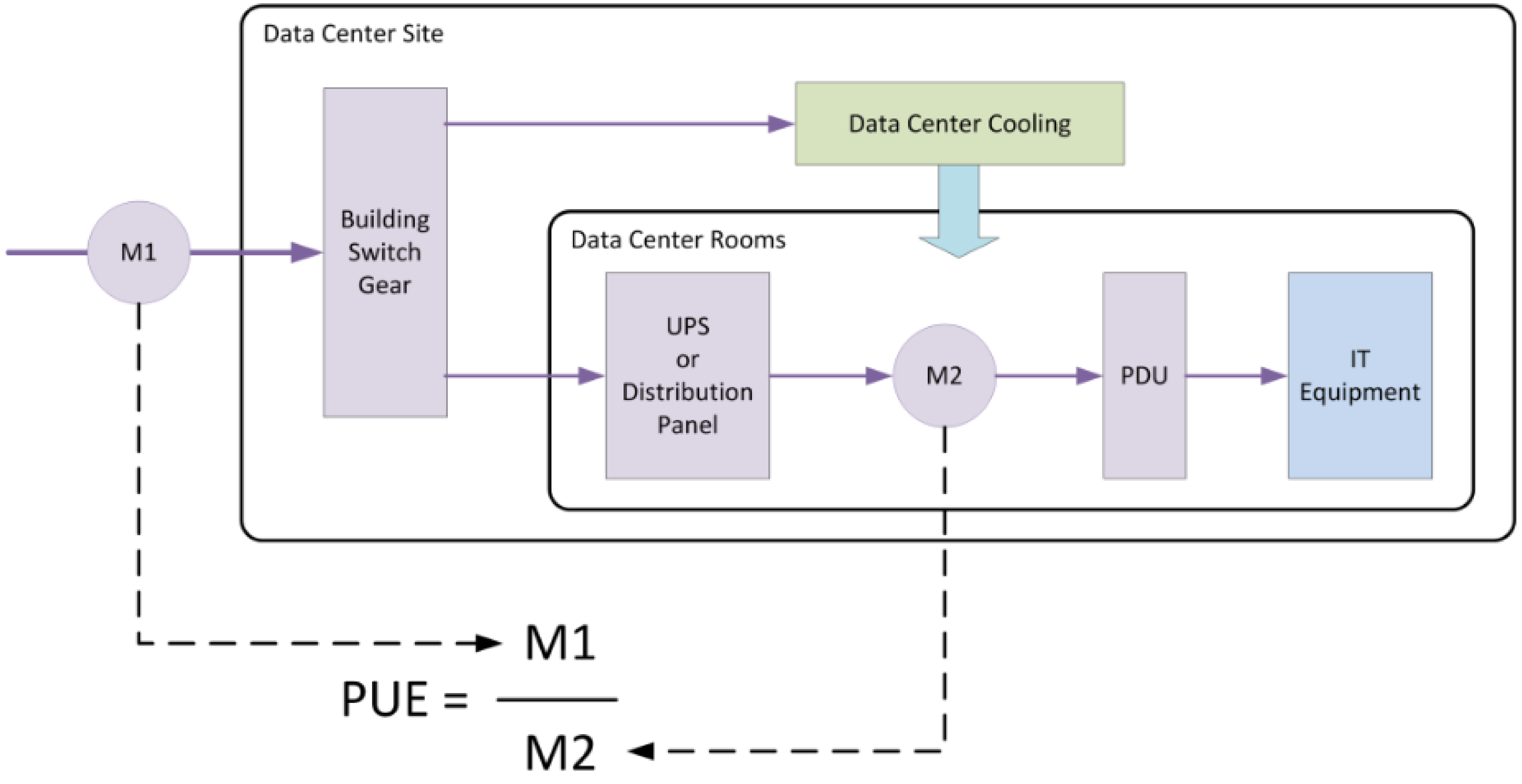
Green500 Data

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)
1	293	**Henri** - ThinkSystem SR670 V2, Intel Xeon Platinum 8362 32C 2.8GHz, NVIDIA H100 80GB PCIe, Infiniband HDR, Lenovo Flatiron Institute United States	8,288	2.88	44	65.396
2	44	**Frontier TDS** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	120,832	19.20	309	62.684

PUE

$$PUE = \frac{Total\ DC\ facility\ energy\ usage}{IT\ equipment\ energy\ usage}$$

- Power Usage Efficiency
- *Time window?*

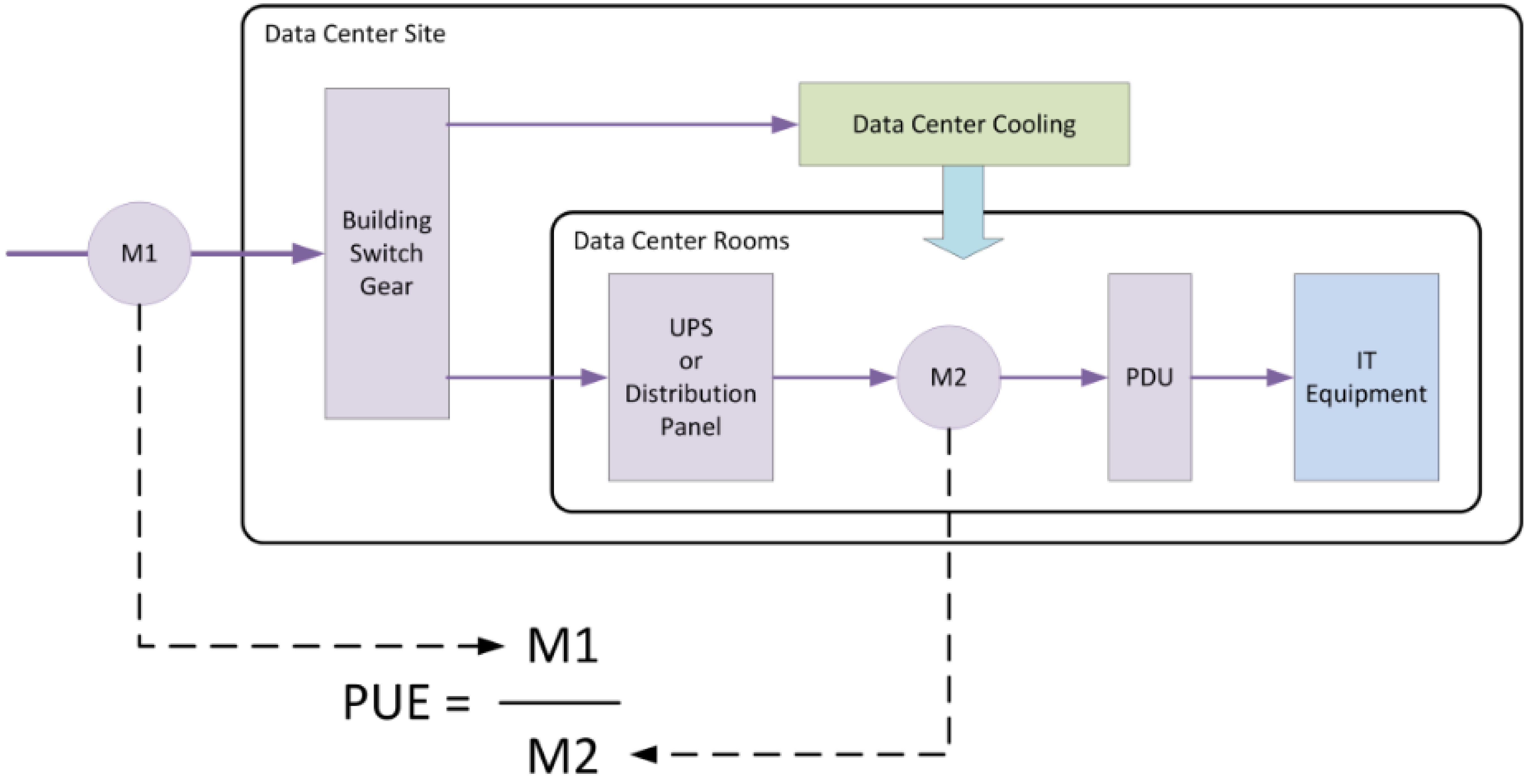


source, Better Buildings, U.S. Department of Energy
https://betterbuildingssolutioncenter.energy.gov/sites/default/files/Get_On_Board.pdf

PUE

$$PUE = \frac{Total\ DC\ facility\ energy\ usage}{IT\ equipment\ energy\ usage}$$

- Power Usage Efficiency
- *Time window?*
- Trailing Twelve-Month (TTM) PUE

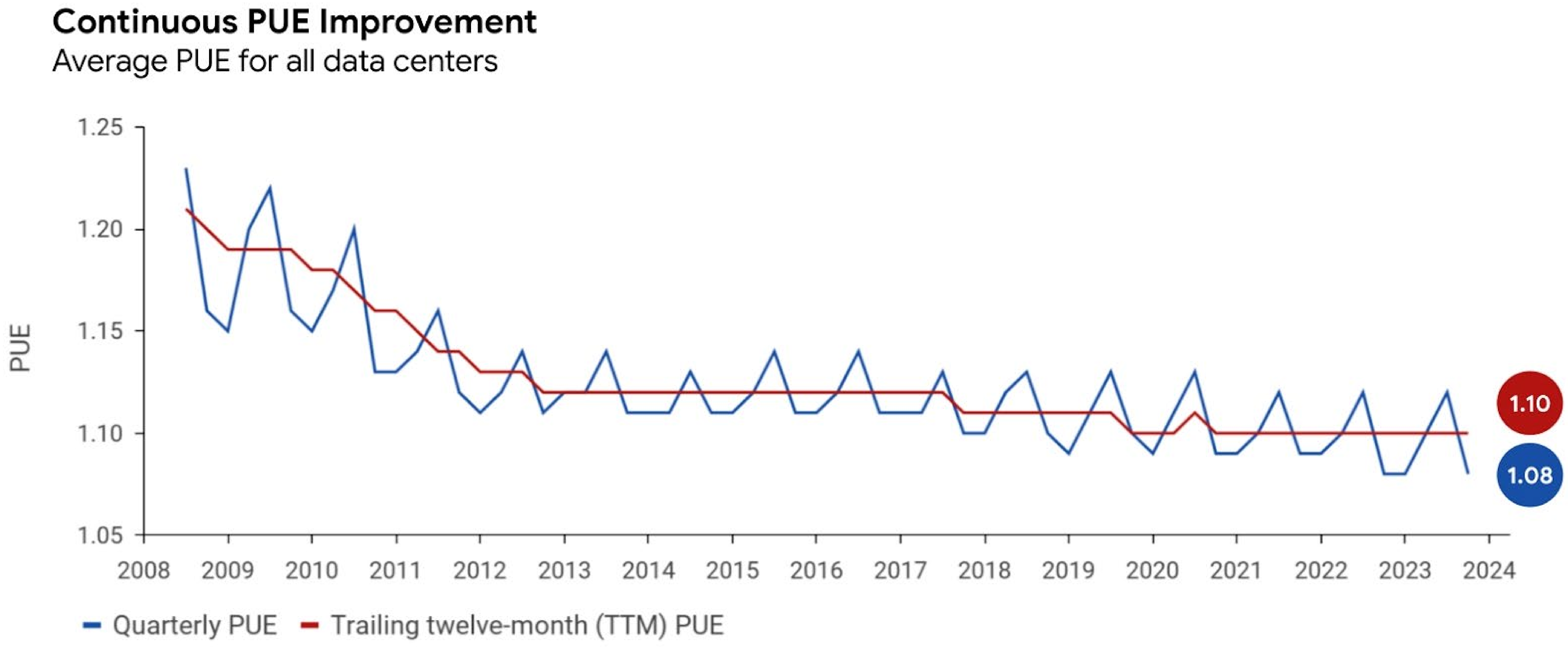


source, Better Buildings, U.S. Department of Energy
https://betterbuildingssolutioncenter.energy.gov/sites/default/files/Get_On_Board.pdf

PUE

$$PUE = \frac{Total\ DC\ facility\ energy\ usage}{IT\ equipment\ energy\ usage}$$

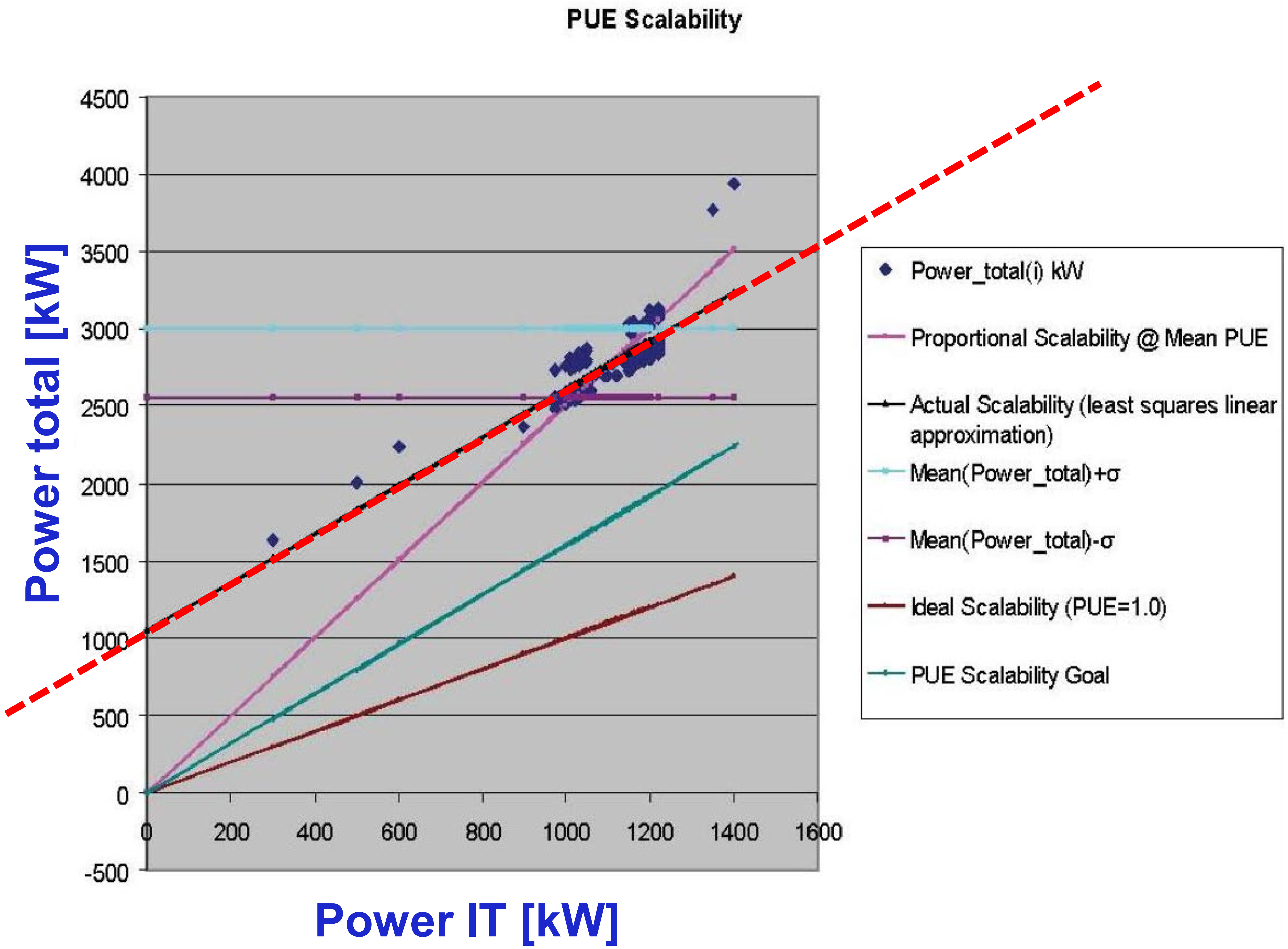
- PUE data for all large-scale Google Data Centers



https://www.google.com/about/datacenters/efficiency/#measuring-efficiency

PUE scalability

- Load
- External conditions



https://datacenters.lbl.gov/sites/default/files/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric_v6.pdf

DC efficiency is not just the energy

- Energy
  - CO_2 emissions – renewable energy sources (sustainablility)
  - Green Energy Consumed (GEC)
- Cooling efficiency
  - Heat re-use
  - Free cooling
- Water

PUE family

- ERE = Energy Reuse Effectiveness

$$ERE = \frac{total\ DC\ facility\ energy\ usage - reuse\ energy}{IT\ equipment\ energy\ usage}$$

PUE family

- ERE = Energy Reuse Effectiveness

$$ERE = \frac{total\ DC\ facility\ energy\ usage - reuse\ energy}{IT\ equipment\ energy\ usage}$$

- ITUE = IT Usage Effectiveness

$$ITUE = \frac{IT\ equipment\ energy\ usage}{compute\ components\ energy\ usage}$$

PUE family

- ERE = Energy Reuse Effectiveness

$$ERE = \frac{total\ DC\ facility\ energy\ usage - reuse\ energy}{IT\ equipment\ energy\ usage}$$

- ITUE = IT Usage Effectiveness

$$ITUE = \frac{IT\ equipment\ energy\ usage}{compute\ components\ energy\ usage}$$

- CUE = Carbon Usage Effectiveness

$$CUE = \frac{total\ CO_2\ emissions\ caused\ by\ total\ DC\ facility\ energy\ usage}{IT\ equipment\ energy\ usage}$$

PUE family

- ERE = Energy Reuse Effectiveness

$$ERE = \frac{total\ DC\ facility\ energy\ usage - reuse\ energy}{IT\ equipment\ energy\ usage}$$

- ITUE = IT Usage Effectiveness

$$ITUE = \frac{IT\ equipment\ energy\ usage}{compute\ components\ energy\ usage}$$

- CUE = Carbon Usage Effectiveness

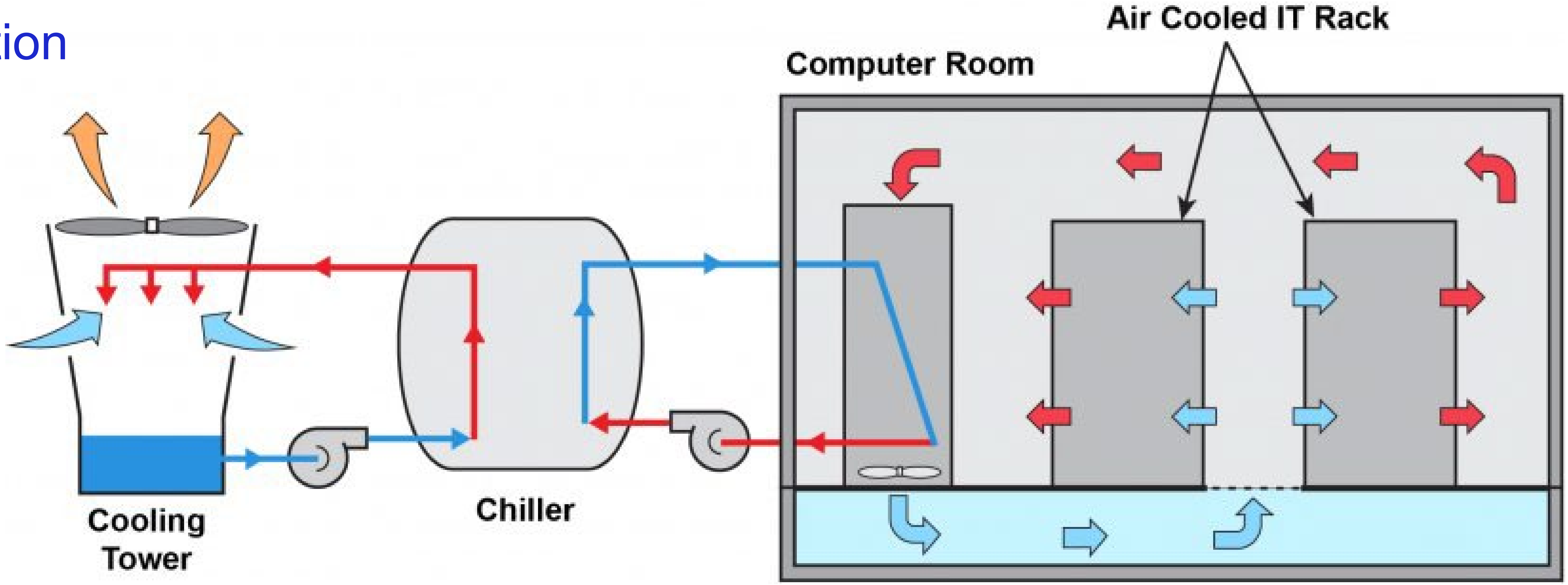
$$CUE = \frac{total\ CO_2\ emissions\ caused\ by\ total\ DC\ facility\ energy\ usage}{IT\ equipment\ energy\ usage}$$

- WUE = Water Usage Effectiveness

$$WUE = \frac{annual\ site\ water\ usage}{IT\ equipment\ energy\ usage}$$

Water savings

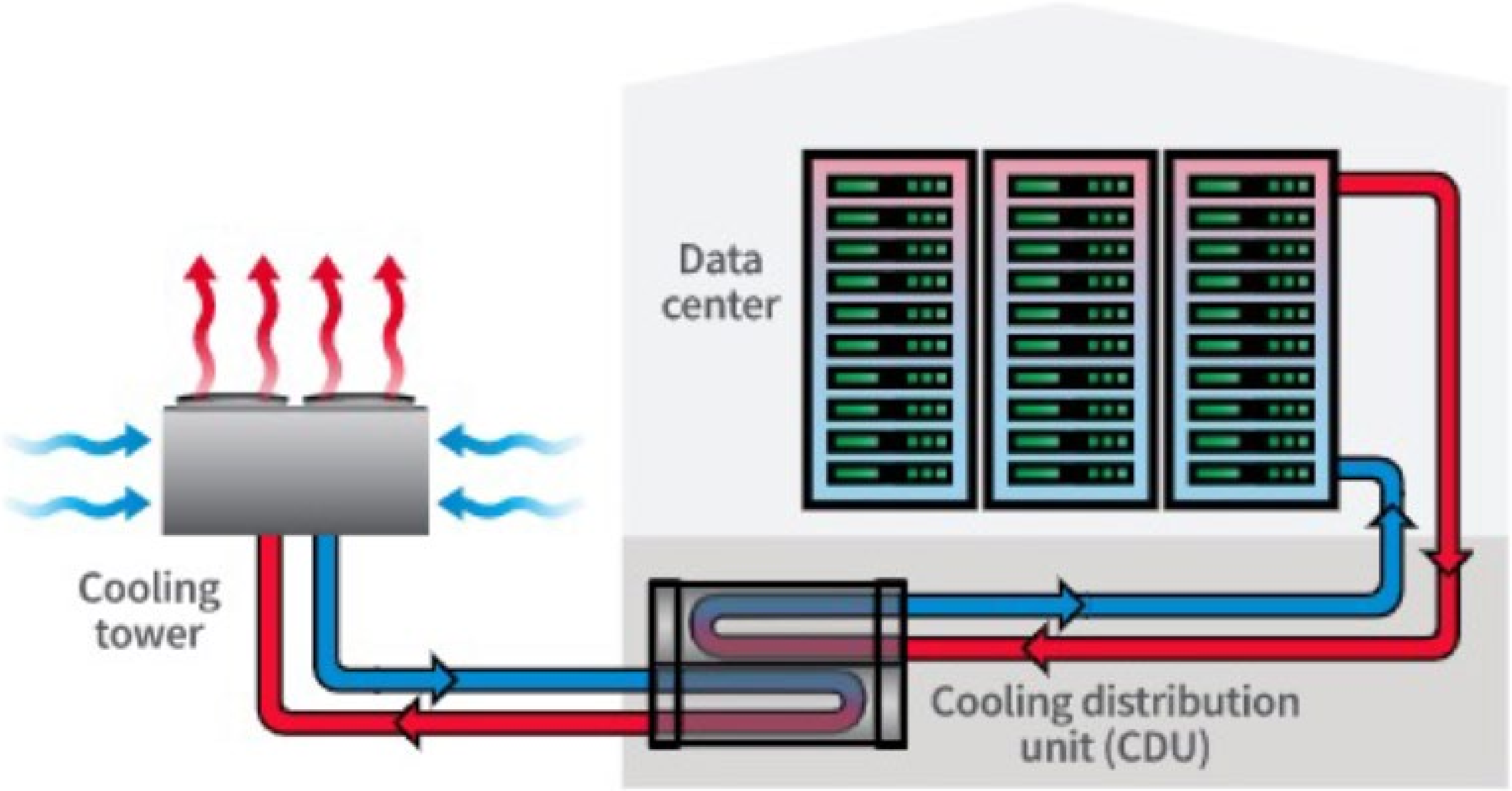
- Space temperature and humidity control
- Air management
- Maximizing Cycles of concetration
- Thermal storage
- Reverse osmosis

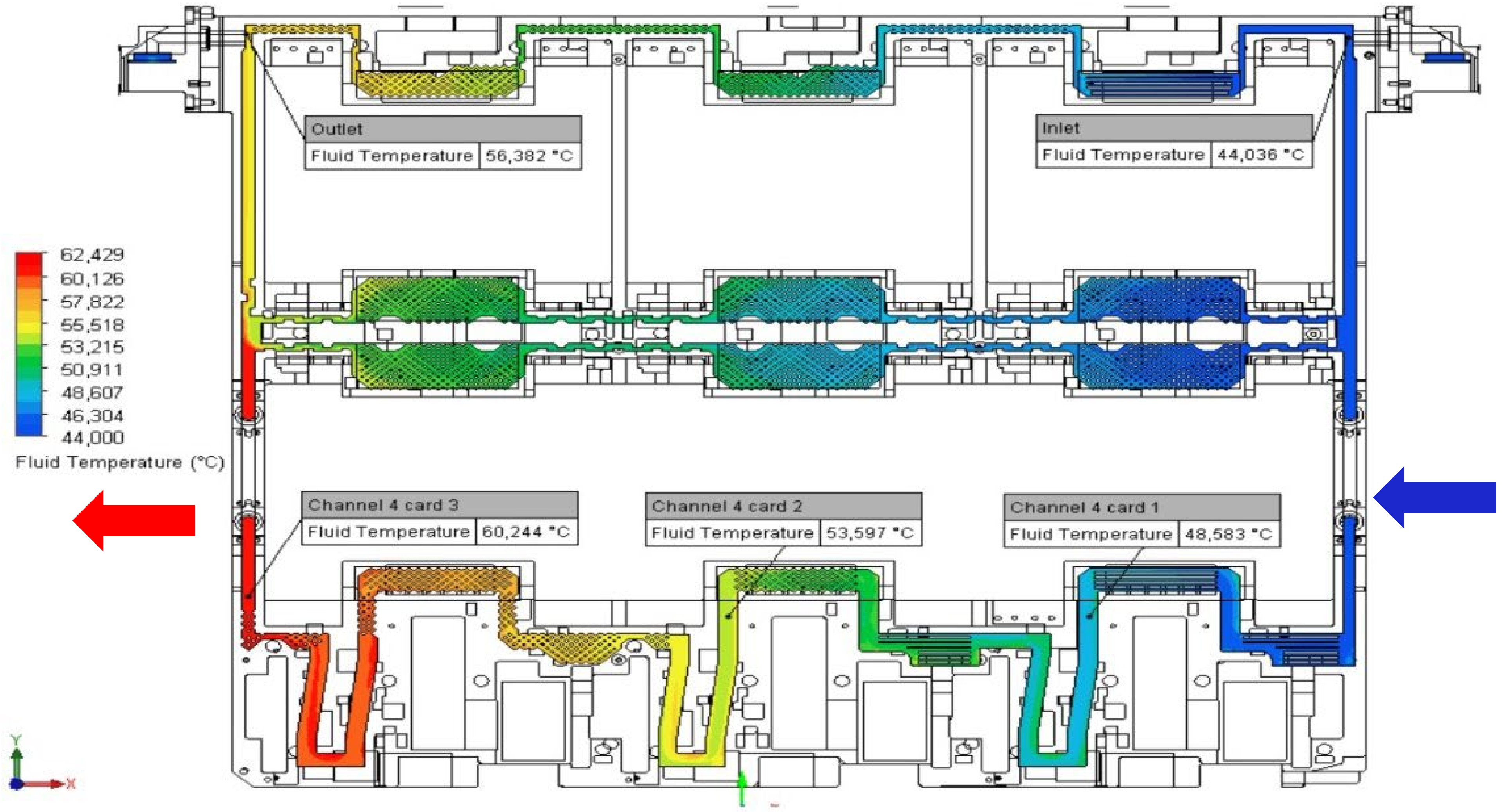


https://www.energy.gov/femp/cooling-water-efficiency-opportunities-federal-data-centers

Warm water direct liquid cooling

- Liquids require significantly less energy
- Liquid cooling - Thermal stability, space efficiency
- Warm water cooling
    - Eliminates inefficient and expensive chillers
    - Eliminates condensation concerns
    - Better waste heat re-use options
    - Save wasted fan energy and use it for computing
    - Prevents thermal capping





Img source Atos, and Supermicro

EuroHPC Karolina

- From 1.2. 2023

$$PUE = \frac{Total\ DC\ facility\ energy\ usage}{IT\ equipment\ energy\ usage}$$

- ~0,06 PUE increase

PUE alternatives

the green grid™

- CPE = compute power efficiency

$$CPE = \frac{IT\ equipment\ utilization * IT\ equipment\ power}{Total\ DC\ facility\ power}$$

- DCeP = Data Center Energy Productivity

$$DCeP = \frac{useful\ work\ produced}{total\ DC\ energy\ consumed\ producing\ this\ work}$$

Jean-Olivier Gerphagnon

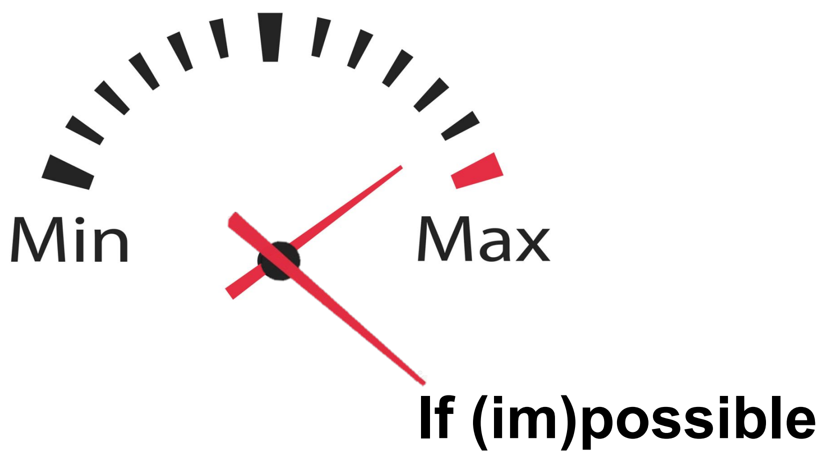
Eviden

Objectives
=

Reality
=

Objectives

=



Reality

=

BUT

BUT

What to measure?

How to measure?

For which purpose?

Easy to compare?

BUT

What to measure?

How to measure?

For which purpose?

Easy to compare?

However…
what is « efficiency »?

Steps

Gathering

Data

Steps

Gathering

Data

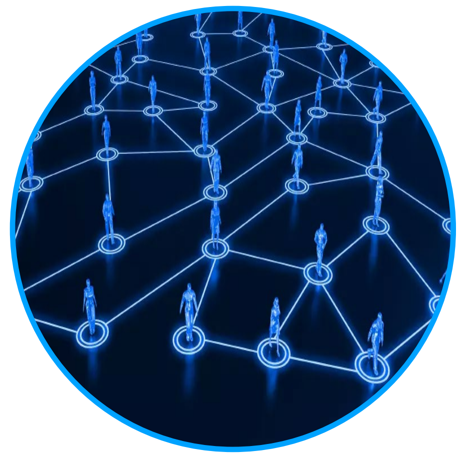
= Monitoring

Steps

Gathering

Data

Modeling



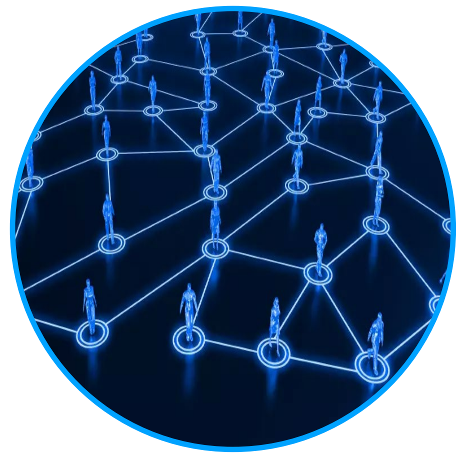
Data

Steps

Gathering

Data

Modeling

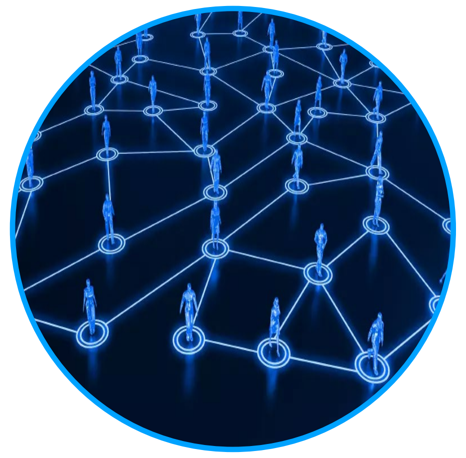


Data

Understanding

Data

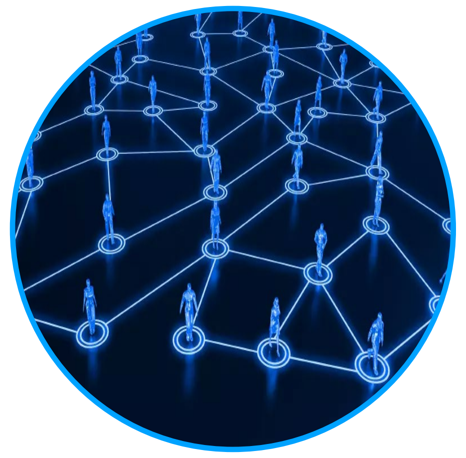
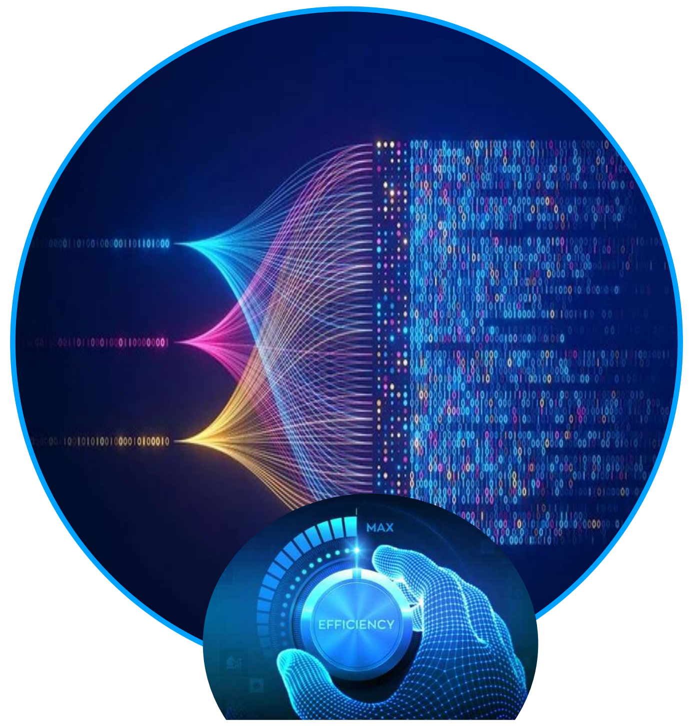
Steps

Gathering	Modeling	Understanding	Using
			
Data	Data	Data	Data

How does One Define An ‘Energy Efficient’ HPC System?

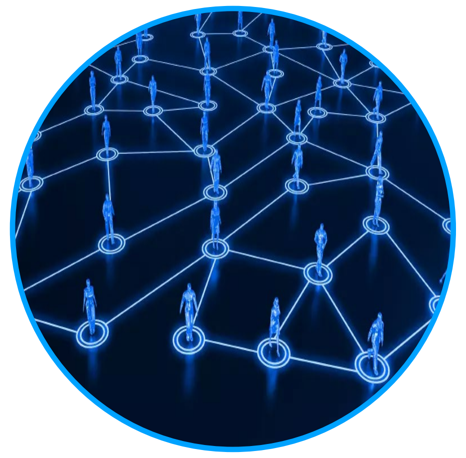
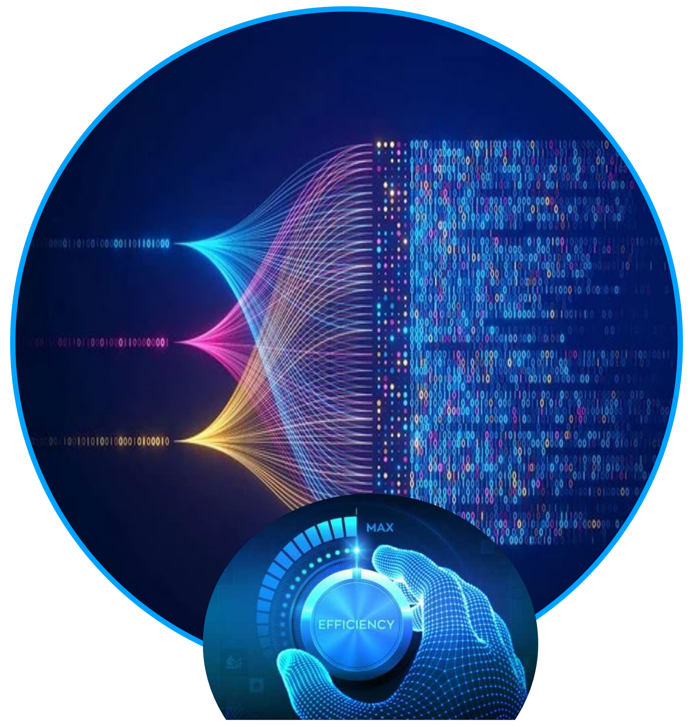
From Data to Action

Steps

Gathering	Modeling	Understanding	Using	Improving
				
Data	Data	Data	Data	**Efficiency**

How does One Define An 'Energy Efficient' HPC System?
From Data to Action

Steps

Gathering	Modeling	Understanding	Using	Improving
				
Data	Data	Data	Data	**Efficiency**

Data

Action

Steps

Gathering	Modeling	Understanding	Using	Improving

Data	Data	Data	Data	Efficiency

Metrics

Energy

Logs

Power

Performance

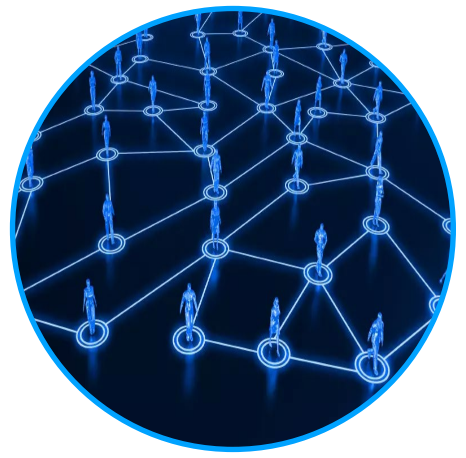
Results

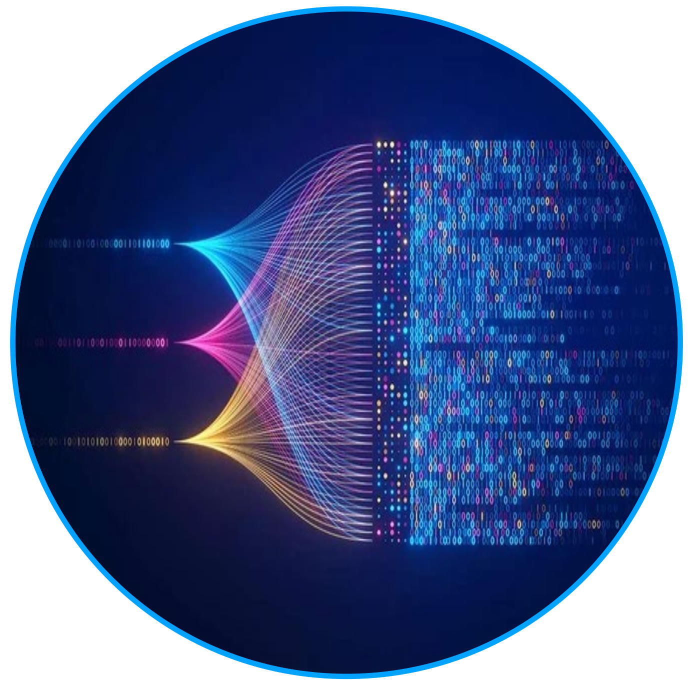
Hardware Counters

Metrics

Events

Steps

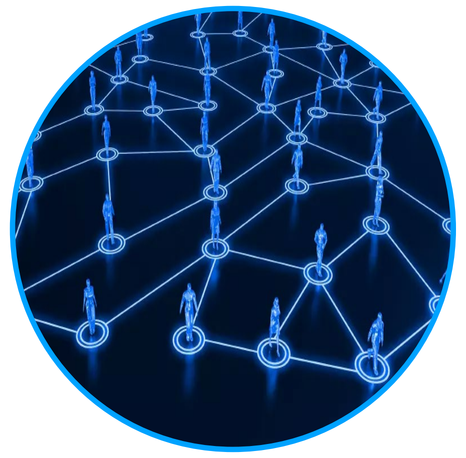
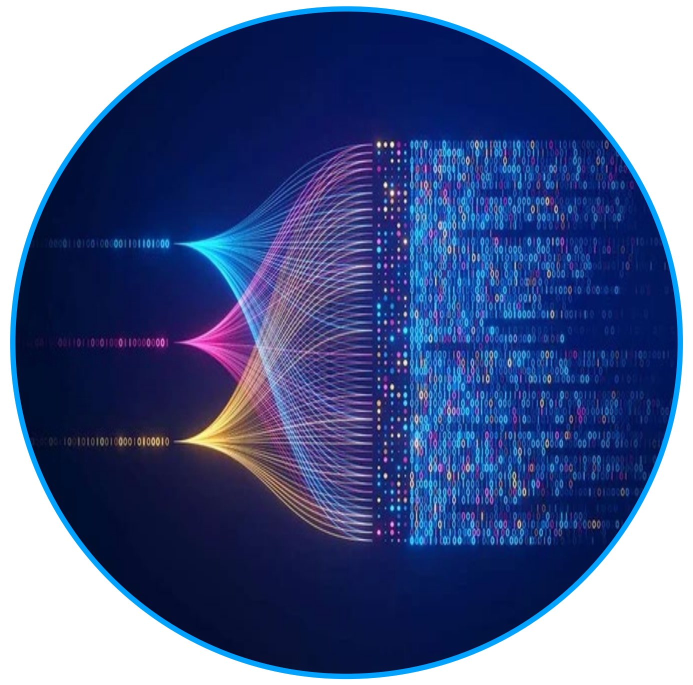



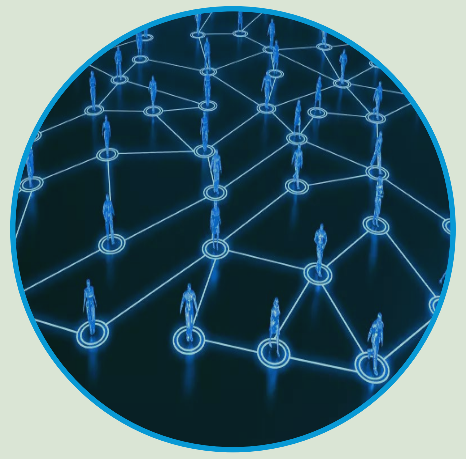
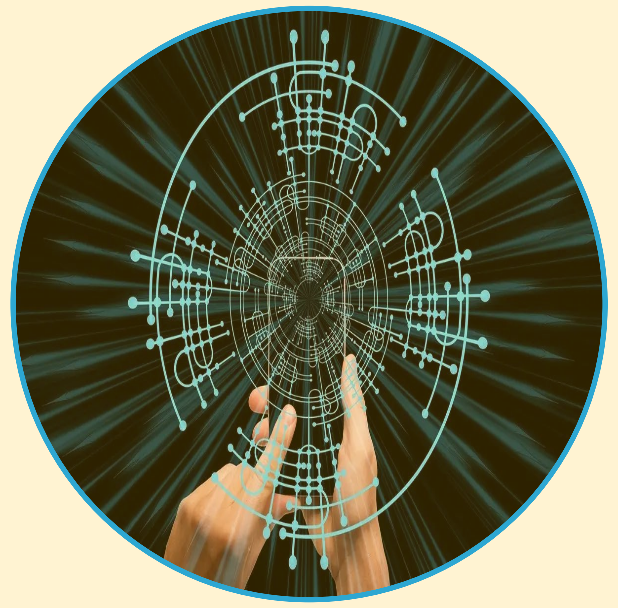
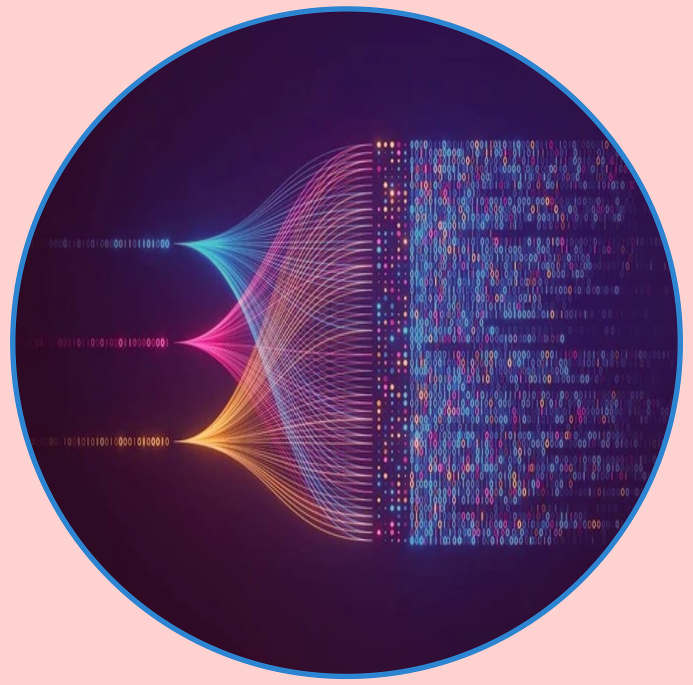
Challenges

	Gathering	Modeling	Understanding	Using	Improving
	Data	Data	Data	Data	Efficiency
Time					
Complexity					
Reproducibility					
HW/SW Dependencies					

How does One Define An 'Energy Efficient' HPC System?
From Data to Action

Steps	Gathering	Modeling	Understanding	Using	Improving
					
	Data	Data	Data	Data	Efficiency
Challenges					
Time	Medium	Low to Medium	High	High	?
Complexity	Low	Low	Medium	High	Very High
Reproducibility	High	High	Medium	?	?
HW/SW Dependencies	Medium	Low	?	High	High

How does One Define An 'Energy Efficient' HPC System?
From Data to Action

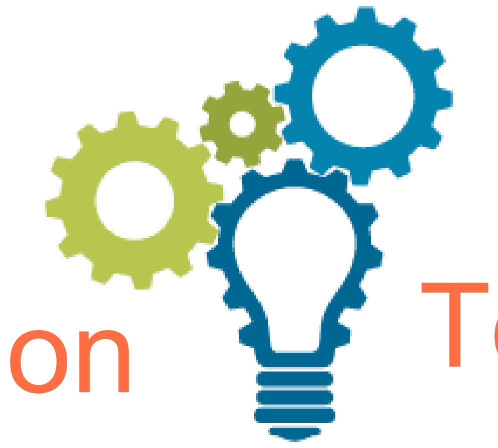
Steps	Gathering	Modeling	Understanding	Using	Improving
					
	Data	Data	Data	Data	Efficiency
Challenges					
Time	Medium	Low to Medium	High	High	?
Complexity	Low	Low	Medium	High	Very High
Reproducibility	High	High	Medium	?	?
HW/SW Dependencies	Medium	Low	?	High	High

Where a provider may be involved?

Where a provider may be involved?
Why

EVIDEN
(or any other ☺)

= Design



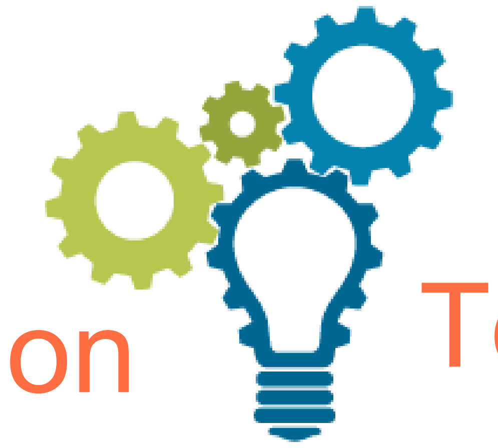
Architecture
Solution
Tools
Standards
Interfaces

Why a provider may be involved?

EVIDEN

(or any other ☺)

= Design



Architecture

Solution

Tools

Standards

Interfaces

Requirements have to be specified in tenders

Hardware + Software + Service = Solution

Why a provider may be involved? EVIDEN

(or any other ☺)

	Hardware	+ Software	+ Service
Env. Footprint			
Energy Efficiency			
Reliability			
Monitor			
Re-use (inc. standards)			

Why a provider may be involved? EVIDEN

(or any other ☺)

	Hardware +	Software +	Service
Env. Footprint	Components type & origin; Transports & Packaging; Easy to replace; Reduced number of components		
Energy Efficiency	Heat dissipation; Water usage; Heat re-use; Power convertor; Efficient cooling; Power control		
Reliability	Resiliency; High MTBF; Quality control		
Monitor	Sensors (in/out band); Frequency; Quality; Interfaces		
Re-use (inc. standards)	Standard slots; Next-gen ready; Capacity enhancement (cooling, power, etc.)		

Why a provider may be involved? EVIDEN (or any other ☺)

	Hardware +	Software +	Service
Env. Footprint	Components type & origin Transports & Packaging Easy to replace Reduced number of components	Environmental aware DevOps Run less model Efficient software (code analysis, etc.)	
Energy Efficiency	Heat dissipation Water usage Power convertor Heat re-use Efficient cooling Power control	Run less model @Runtime live optimization Monitor Data	
Reliability	Resiliency High MTBF Quality control	Validation Native HA Up-to-date Code Analysis Live management	
Monitor	Sensors (in/out band) Frequency Quality Interfaces Granularity	Non-impacting data gathering Infra-oriented Historical Job-oriented	
Re-use (inc. standards)	Standard slots Next-gen ready Capacity enhancement (cooling, power, etc.)	Open-source based Rely on APIs Open-access model EU projects	

Why a provider may be involved? EVIDEN

(or any other ☺)

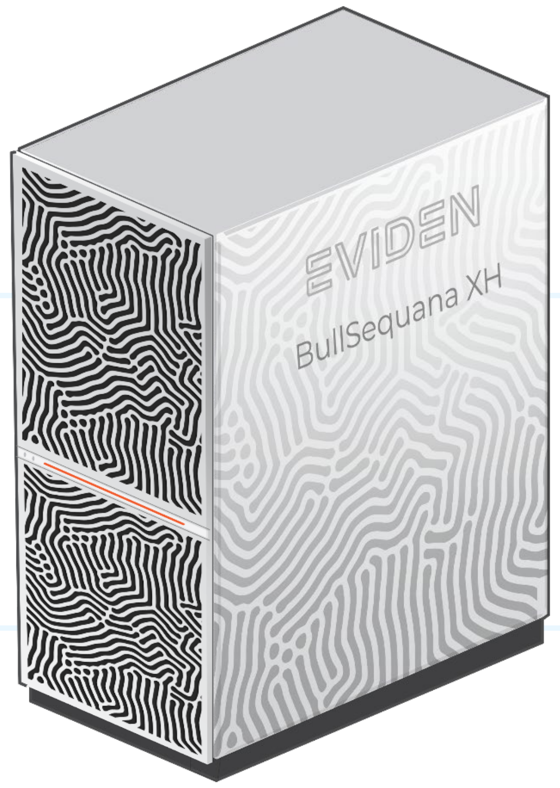
	Hardware +	Software +	Service
Env. Footprint	Components type & origin Transports & Packaging Easy to replace Reduced number of components	Environmental aware DevOps Efficient software (code analysis, etc.) Run less model	
Energy Efficiency	Heat dissipation Water usage Heat re-use Power convertor Efficient cooling Power control	Run less model @Runtime live optimization Monitor Data	Expertise
Reliability	Resiliency High MTBF Quality control	Validation Native HA Up-to-date Code Analysis Live management	Integration
Monitor	Sensors (in/out band) Frequency Quality Interfaces Granularity	Non-impacting data gathering Infra-oriented Historical Job-oriented	
Re-use (inc. standards)	Standard slots Next-gen ready Capacity enhancement (cooling, power, etc.)	Open-source based Rely on APIs Open-access model EU projects	

Why a provider may be involved? EVIDEN

(or any other ☺)

	Hardware +	Software +	Service =
Env. Footprint	Components type & origin Transports & Packaging Easy to replace Reduced number of components	Environmental aware DevOps Efficient software (code analysis, etc.) Run less model	Expertise Integration
Energy Efficiency	Heat dissipation Water usage Heat re-use Power convertor Efficient cooling Power control	Run less model @Runtime live optimization Monitor Data	Expertise Integration
Reliability	Resiliency High MTBF Quality control	Validation Native HA Up-to-date Code Analysis Live management	Expertise Integration
Monitor	Sensors (in/out band) Frequency Granularity Quality Interfaces	Non-impacting data gathering Historical Infra-oriented Job-oriented	Expertise Integration
Re-use (inc. standards)	Standard slots Next-gen ready Capacity enhancement (cooling, power, etc.)	Open-source based Rely on APIs Open-access model EU projects	Expertise Integration

Solution



SMC xScale

SEMS
Smart Energy Management Suite

Performance Studio

Why we must all be involved?

Hardware + Software + Service = Solution

Env. Footprint

Energy Efficiency

Reliability

Monitor

Re-use (inc. standards)

1. Get & structure data (even if you don't know yet why)
2. Share & combine data (users, sites, communities, providers, etc.)
3. Define (steped) targets of improvement
4. Set (realistic) requirements in tenders (and discuss far in advance)

Sensors (in/out band)

Quality

Standard slots

EuroHPC Joint Undertaking

EUROHPC
SUMMIT 2024

ANTWERP
18–21 MARCH

TO EXASCALE
AND BEYOND

Per Öster

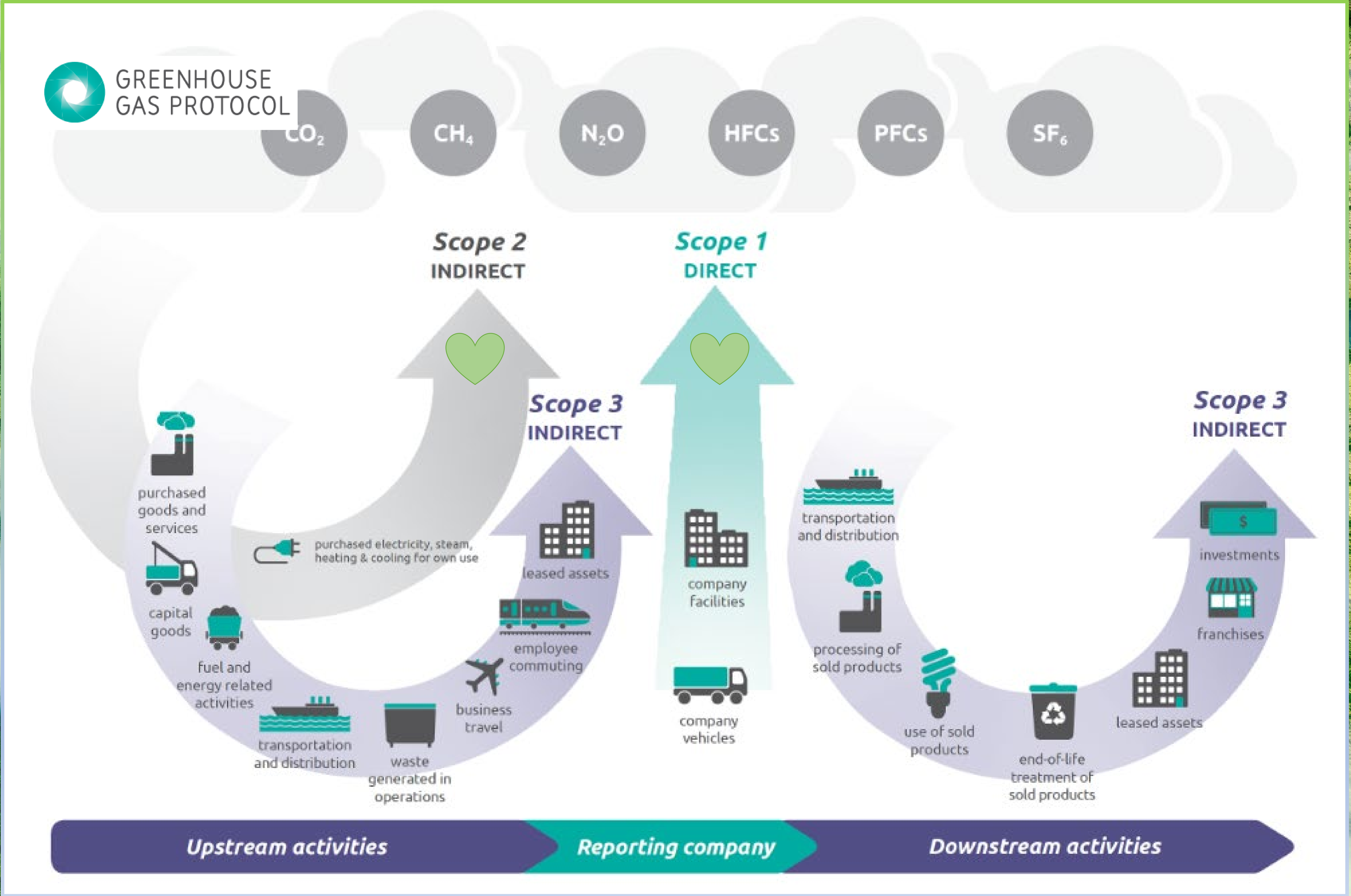
CSC – IT-Center for Science

How Does One Define an "Energy Efficient" HPC System? From Data to Action

ANTWERP

LUMI Data Center, Cooling and Heat Reuse

Per Öster, CSC – IT Center for Science


GREENHOUSE GAS PROTOCOL
CO_2
CH_4
N_2O
HFCs
PFCs
SF_6
Scope 2
INDIRECT
Scope 1
DIRECT
Scope 3
INDIRECT
Scope 3
INDIRECT
purchased goods and services
capital goods
fuel and energy related activities
transportation and distribution
waste generated in operations
business travel
employee commuting
leased assets
purchased electricity, steam, heating & cooling for own use
company facilities
company vehicles
transportation and distribution
processing of sold products
use of sold products
end-of-life treatment of sold products
leased assets
franchises
investments
Upstream activities
Reporting company
Downstream activities

The green home of LUMI: Renforsin Ranta Business Park, Kajaani, Finland

Home of LUMI

Renforsin Ranta Business Park, Kajaani Finland

LUMI is an HPE Cray EX Supercomputer

LUMI

SUSTAINED PERFORMANCE

380 PETAFLOP/S

Since 2022 #1 Supercomputer in Europe

Hewlett Packard Enterprise

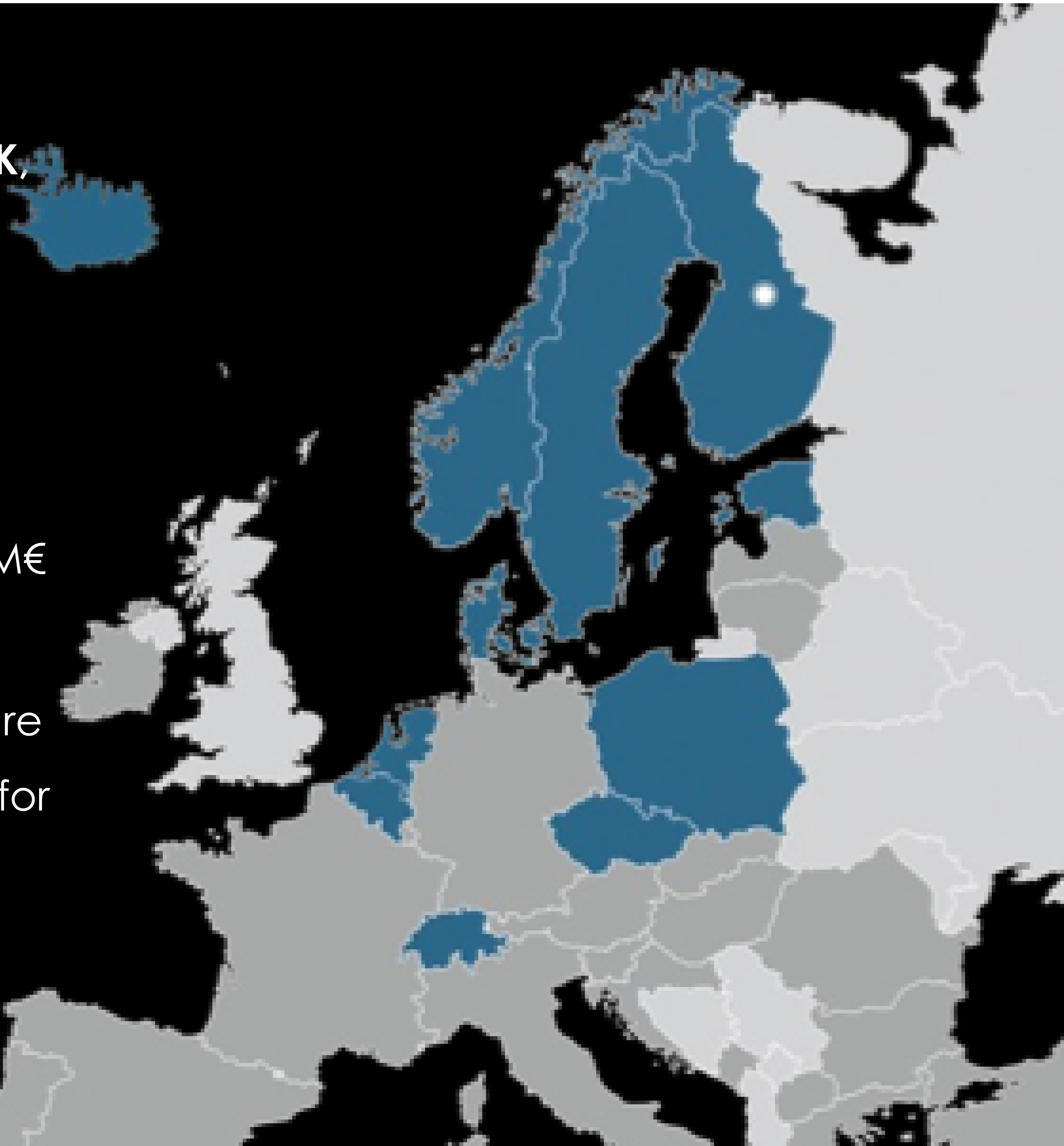
LUMI EuroHPC
Fastest Supercomputer in Europe

LUMI is a co-investment of eleven LUMI consortium countries (FI, BE, CH, CZ, DK, EE, IS, NL, NO, PL, SE) and the EU to build and operate a world-class supercomputer

The EuroHPC Joint Undertaking pools EU and national resources in high-performance computing

The total budget of LUMI in CSC's data center in Kajaani, Finland, is over 202 M€ (50% EU and 50% the consortium countries)

The resources of LUMI are allocated proportionally to the investments. The share of the EuroHPC JU (50%) is allocated by a peer-review process and available for all European researchers



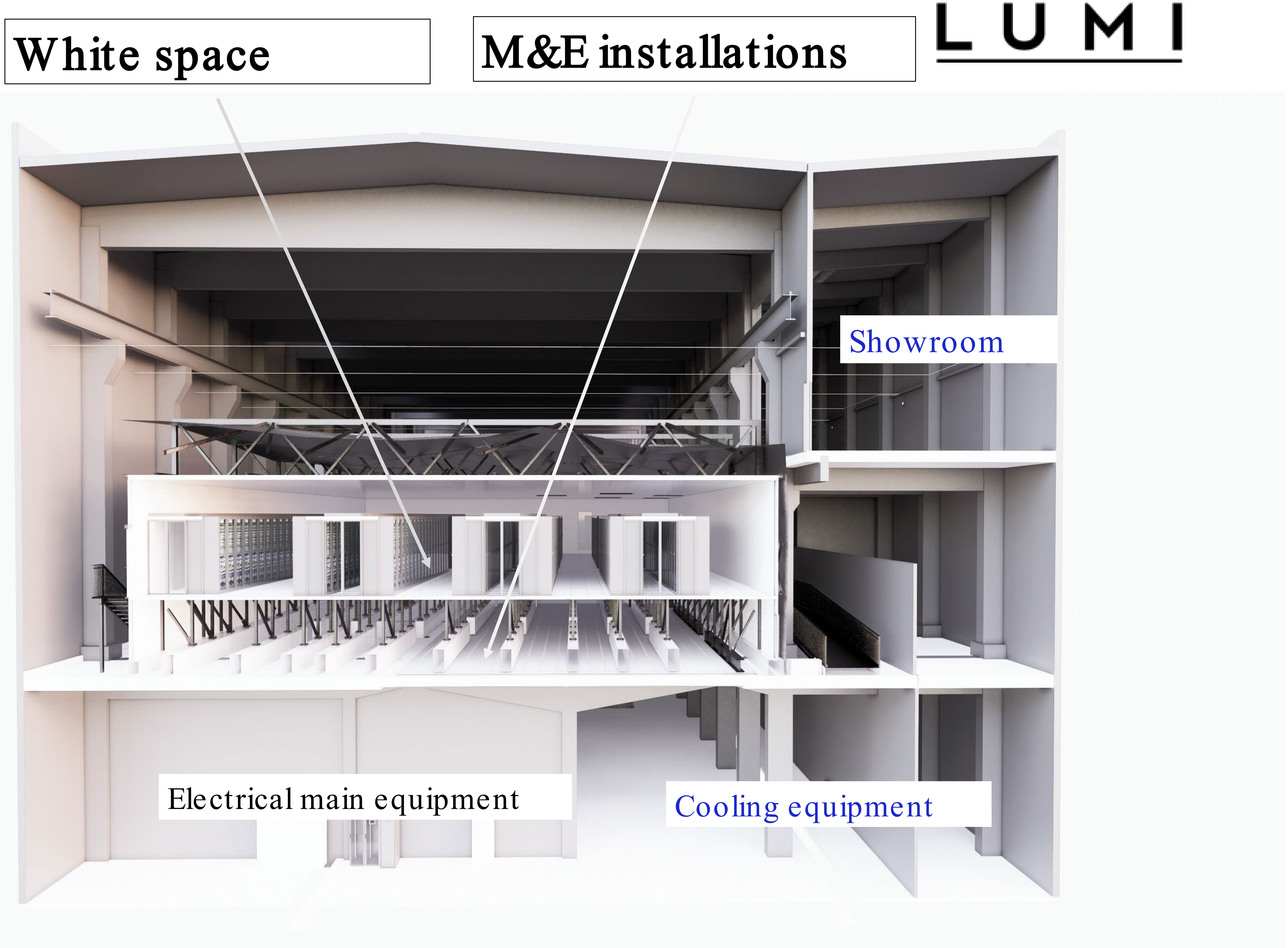
LUMI

LUMI facility overview

- 5800m² in three floors
- 800m² whitespace for IT devices
- Power capacity 15MW at full buildout
- 14 000m² free space for future expansions

Data center cooling

- Mechanical cooling with 3 heat pumps with 7.2MW total cooling capacity and 9MW of total heating capacity
- Free cooling, total capacity 10MW
- Air cooling installed capacity 1MW with N+1 redundancy
- 17°C inlet to CRAH units in whitespace
- Free cooling on chillers >15°C outside temperature



Some Basics of DC Business

LUMI

Electricity consumption:

- **Datacenter capacity (and size)** is primarily measured by electricity consumption in kilowatts (kW) or **megawatts (MW)**
- **Main cost in datacenter operations**

Power Usage Effectiveness (PUE):

- PUE measures the total energy use of the data center compared to the energy used by IT equipment.

Energy Reuse Factor (ERF):

- The ERF of a data center reflects how much energy is exported for reuse outside of data center operations.

Energy Reuse Effectiveness (ERE):

- The ERE is a metric for measuring the benefit of reuse energy from a data center

$$\text{PUE} = \frac{\text{Total Facility Energy (Cooling+Power+Lightning+IT)}}{\text{IT Equipment Energy (IT)}}$$

$$\text{ERF} = \frac{\text{Energy Reuse}}{\text{Total Facility Energy}}$$

$$\text{ERE} = \frac{\text{Total Facility Energy} - \text{Energy Reuse}}{\text{IT Equipment Energy (IT)}}$$

LUMI: Excess Heat Utilization Process Overview

LUMI

EUROHPC SUMMIT 2024

Annual CO_2 savings around 6.000 tonnes

DRY AIR COOLING FOR BACK UP ~ 10 MW

DISTRICT HEATING NETWORKS ~ 10 MW
-Renforsin Ranta Business Park
-CITY of Kajaani

+43 °C

+58 °C

+80 °C

+31 °C

Service demarcation point for the excess heat utilization

HEAT EXCHANGERS

HEAT PUMPS

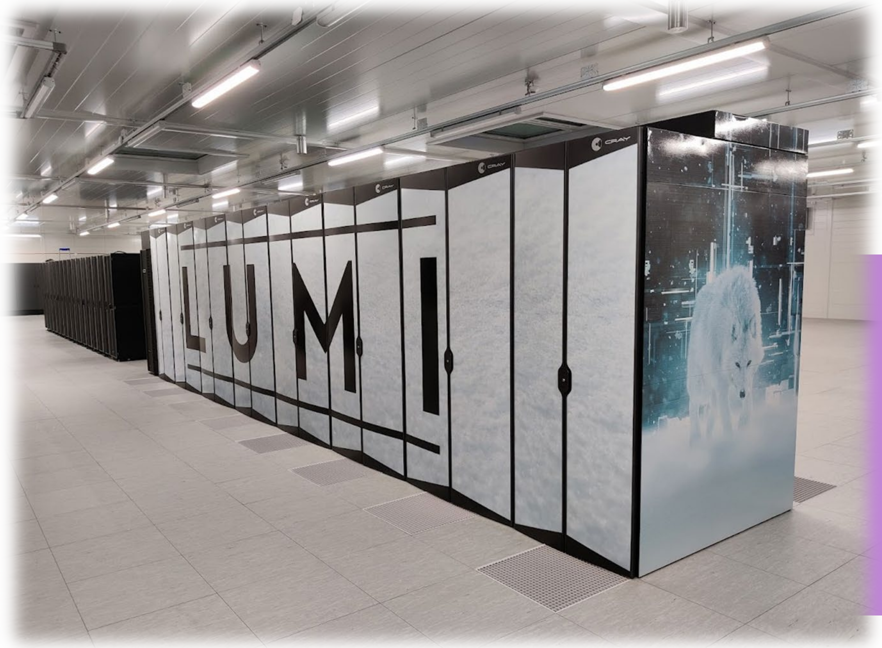
Heat pumps use renewable energy

+32 °C

+44 °C

+32 °C

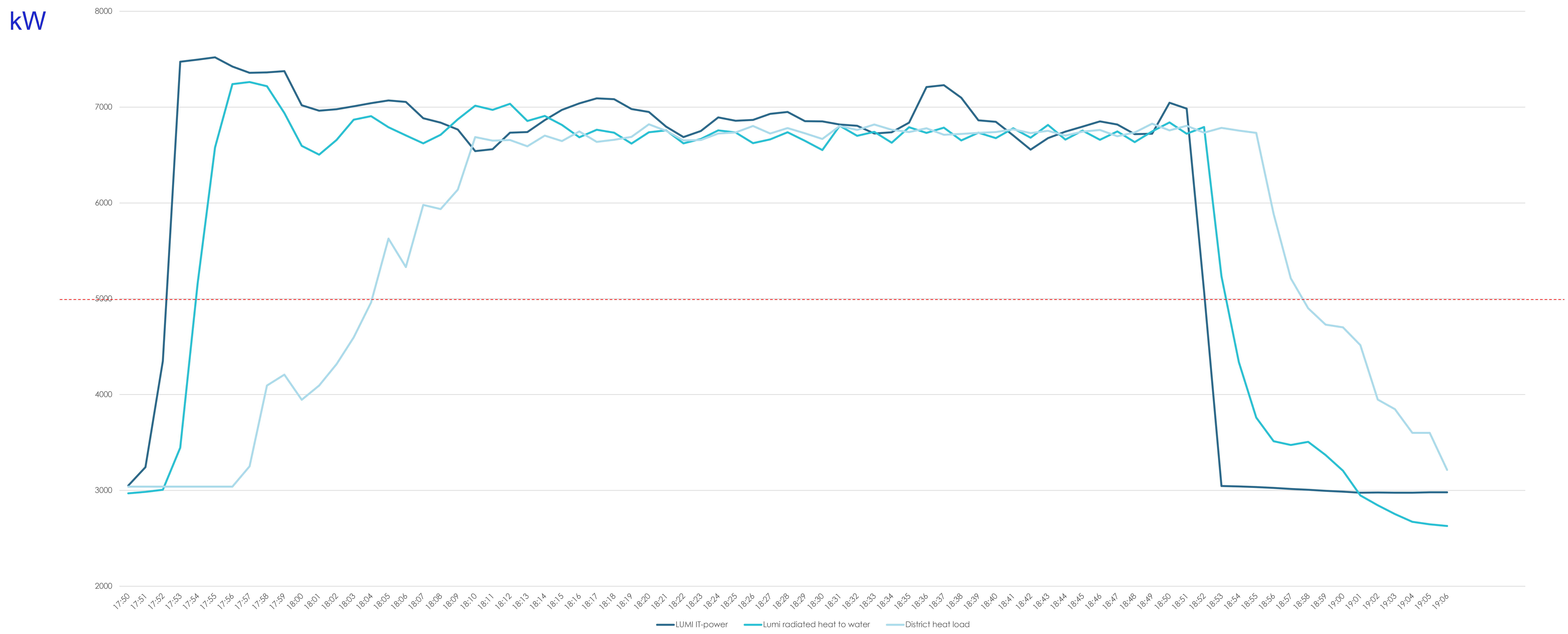
+44 °C



HPC load

In addition of Direct Liquid Cooling there is approximately 1 MW of capacity for the air-cooled servers (e.g. storage and management servers).

Example of power usage and heat rejection delay



LUMI data center sustainability and metrics

- **Waste heat is re-used for district heating ~9 months / year**. Dependent on district heating network demand because power plant has minimum operating power.
- Average temperature in Kajaani in 2022 was 3.6°C and highest measured temperature 2023 was 26.2°C. **Free cooling available around a year.**

- The **brownfield solution** (re-use of exiting building) is estimated to have **reduced the CO_2 footprint of LUMI data center construction by over 80%** compared to constructing an all-new building for LUMI. Approximately 1000 tons.

LUMI

- Changes in Finnish energy taxation for data centers on 2022
- Energy taxation classes for data centers drives more energy efficient data centers with lower Opex.
  - Energy taxation Class 1, 22,53 €/MWh
  - **Energy taxation Class 2, 6,3 €/MWh**
- Data centers whose IT-power exceeds 0,5MW and who are full filling energy efficiency requirements are eligible to apply energy taxation class 2.
- ERE or PUE requirement that must be achieved
  - **ERE**
    - Data center with 0.5..5 MW IT-power calendar year average, ERE < 0.90
    - **Data center with 5..10 MW IT-power calendar year average, ERE < 1.00**
    - Data center over 10 MW IT-power calendar year average, ERE requirement does not apply
  - **PUE**
    - Data center calendar year average PUE < 1.25
    - From 2026 calendar year average PUE < 1.20

LUMI

Metric	LUMI (values from 2023)
Free cooling PUE	<1.05
PUE with heat re-use	1.31
Heat re-use COP	4
ERE	0.52
ERF	0.57 = 57%
Annual heat production	26,7 GWh
Reduced CO_2 emission	~ 6.000 CO_2 tonnes *
Source of electricity	100% hydroelectric power

* LUMI reduced CO_2 emisson based on district heating production plant average CO_2 emission / year (2022: 161kg/MWh). No real-time data available

LUMI waste heat reuse lessons learnt

LUMI

- Heat pumps are difficult to operate on HPC load
    - Applies mostly to GPU partition where load swings are large based on GPU utilization
    - Buffer tank or another mechanism to CDU outlet side to align heat load variation to heat pumps
- Partial waste heat utilization should be implemented if there is more than one partner that uses waste heat
- GPU real operational power load is way lower than TDP values or during HPL (50..75% of HPL load)
- Outlet temperature varies based on the IT load of the HPC system -> hard to maintain optimal circumstances for heat production

- Esa Heiskanen, LUMI Operations Manager
- Tero Tuonen, Director Sustainability and Risk Management

LUMI

Per Öster

Director, Advanced Computing Facility

CSC – IT Center for Science Ltd.

per.oster@csc.fi

+358 50 381 9030

Follow us

X: @LUMIhpc

LinkedIn: LUMI supercomputer

YouTube: LUMI supercomputer

www.lumi-supercomputer.eu

contact@lumi-supercomputer.eu

EuroHPC Joint Undertaking

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC Joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the of Participating States FI, BE, CH, CZ, DK, EE, IS, NO, PL, SE.

Leverage from the EU 2014–2020

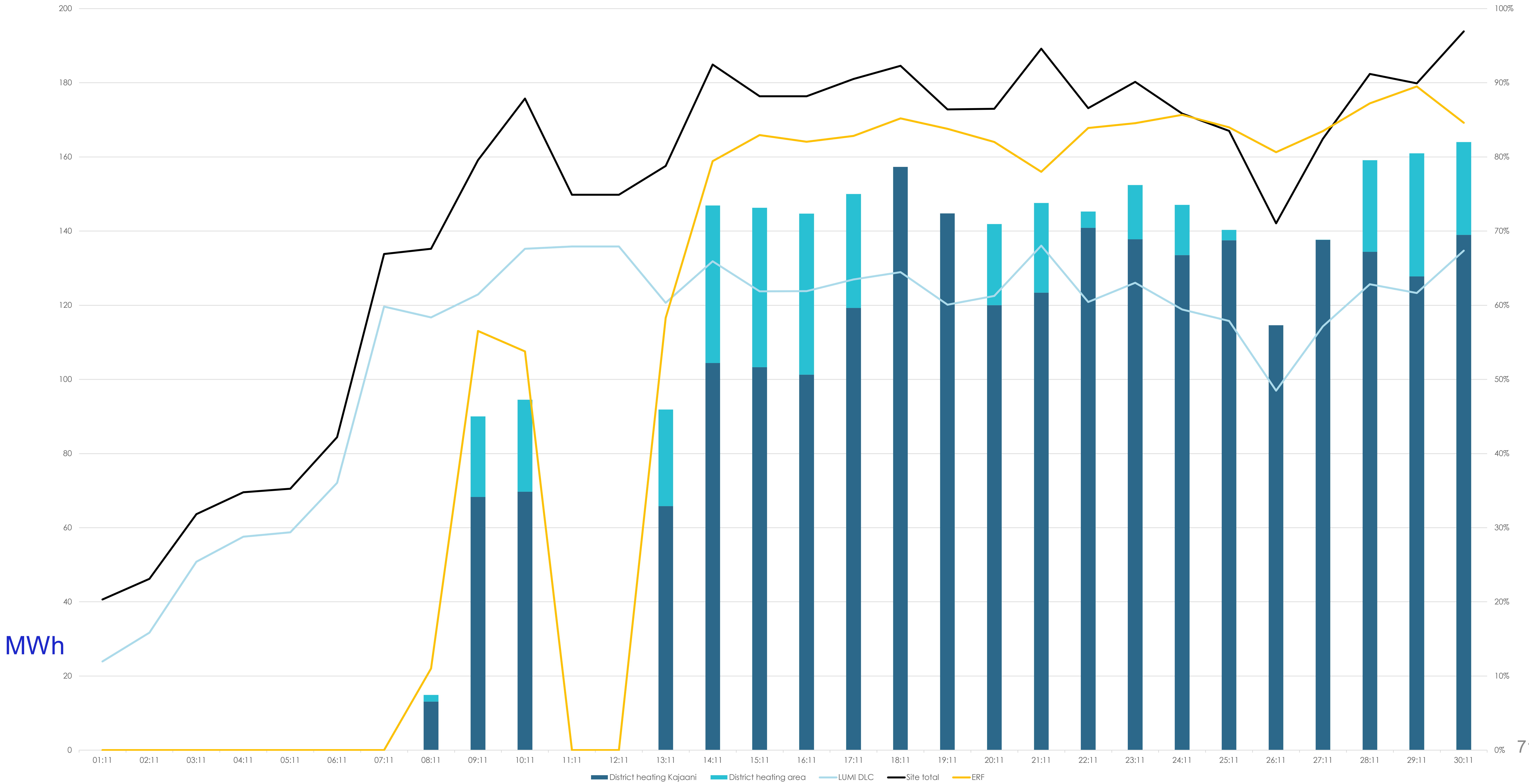
European Union European Regional Development Fund

REGIONAL COUNCIL OF KAINUU

EURO

Additional slides

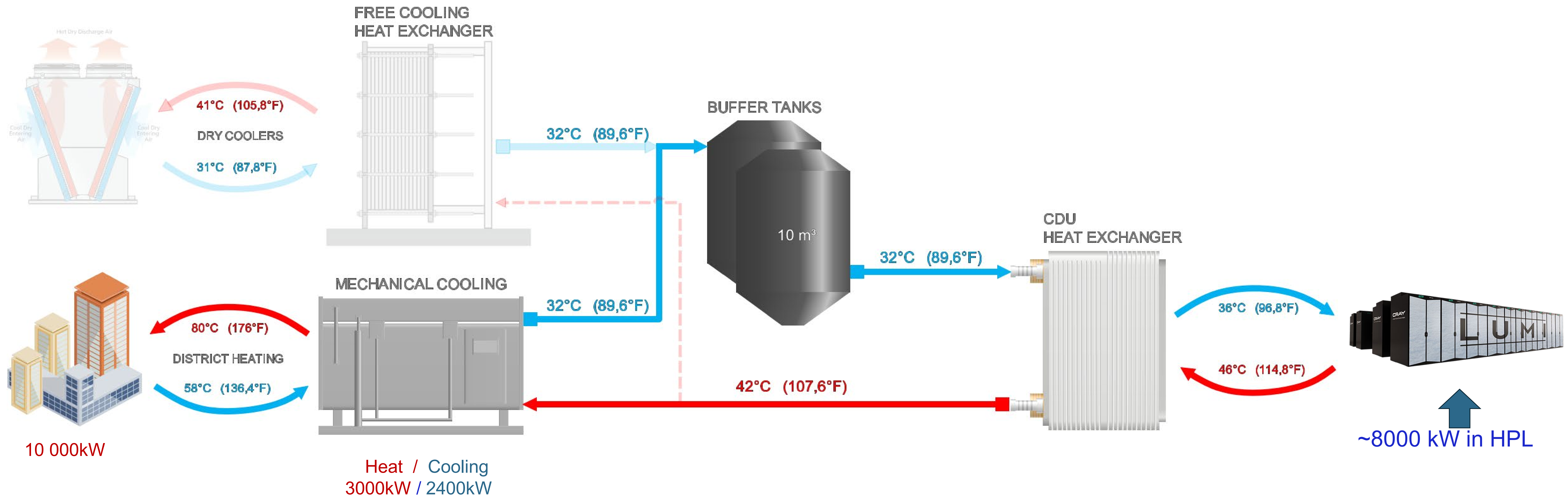
LUMI data center energy consumption and district heating



LUMI

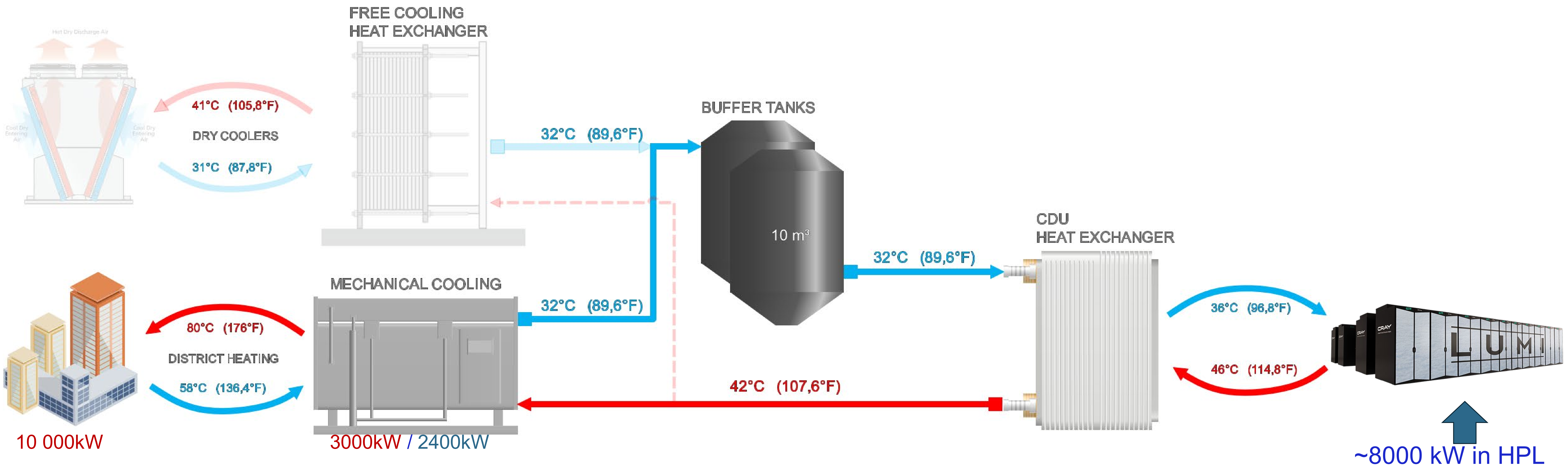
Three individually controlled cooling modules started in 2000kW steps based on LUMI IT power

- **Mechanical cooling** and heat reuse for district heating with three heat pumps.
  - 2400kW cooling and 3000 kW of heating capacity/heat pump.



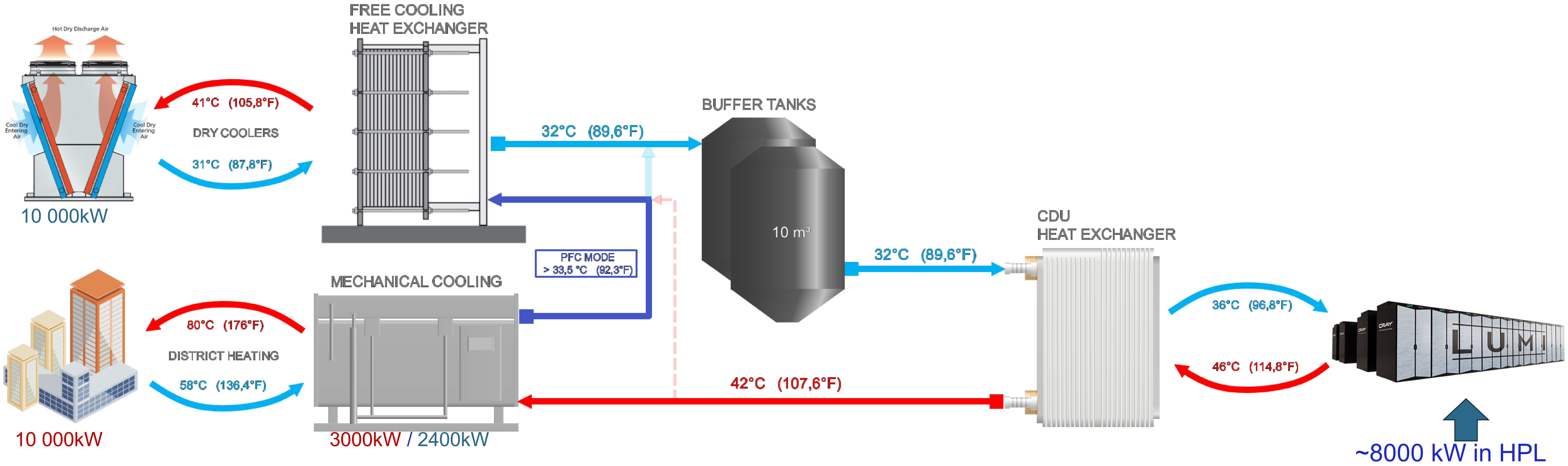
Three individually controlled cooling modules started in 2300kW steps based on LUMI IT power

- **Mechanical cooling** and heat reuse for district heating with three heat pumps.
  - 2400kW cooling and 3000 kW of heating capacity/heat pump.



Three individually controlled cooling modules started in 2300kW steps based on LUMI IT power

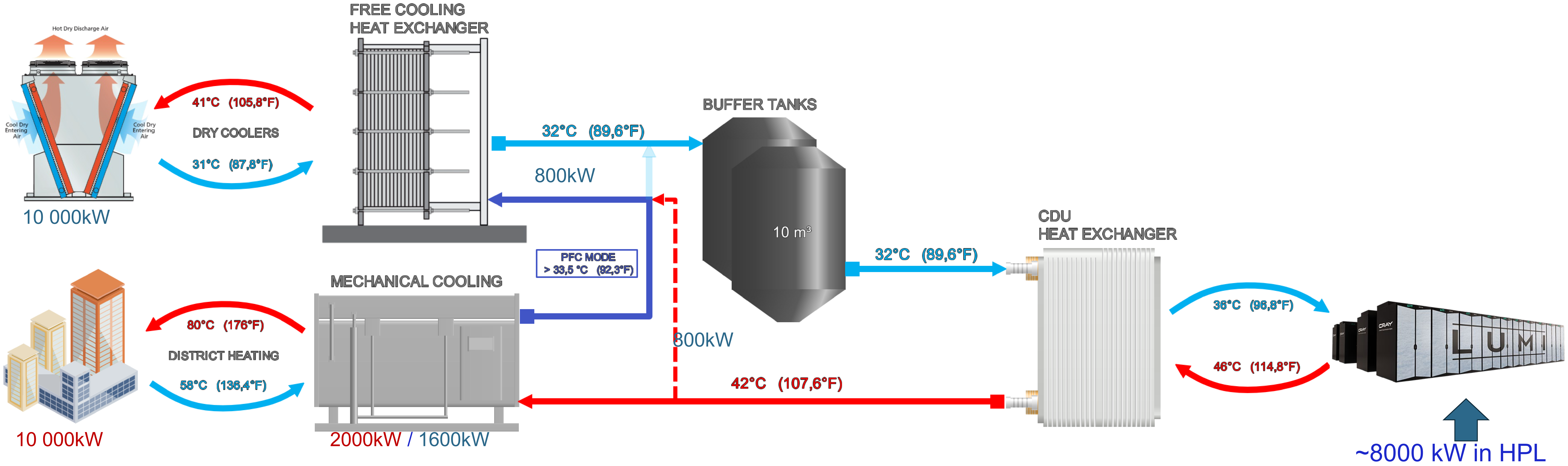
- **Partial free cooling** with heat pumps and free cooling heat exchangers + DCRs.
  - Partial free cooling is used 15 minutes in CGM startup, if there is an alarm on heat pump or the heat pump outlet is >1.5 °C over setpoint



LUMI

Three individually controlled cooling modules started in 2300kW steps based on LUMI IT power

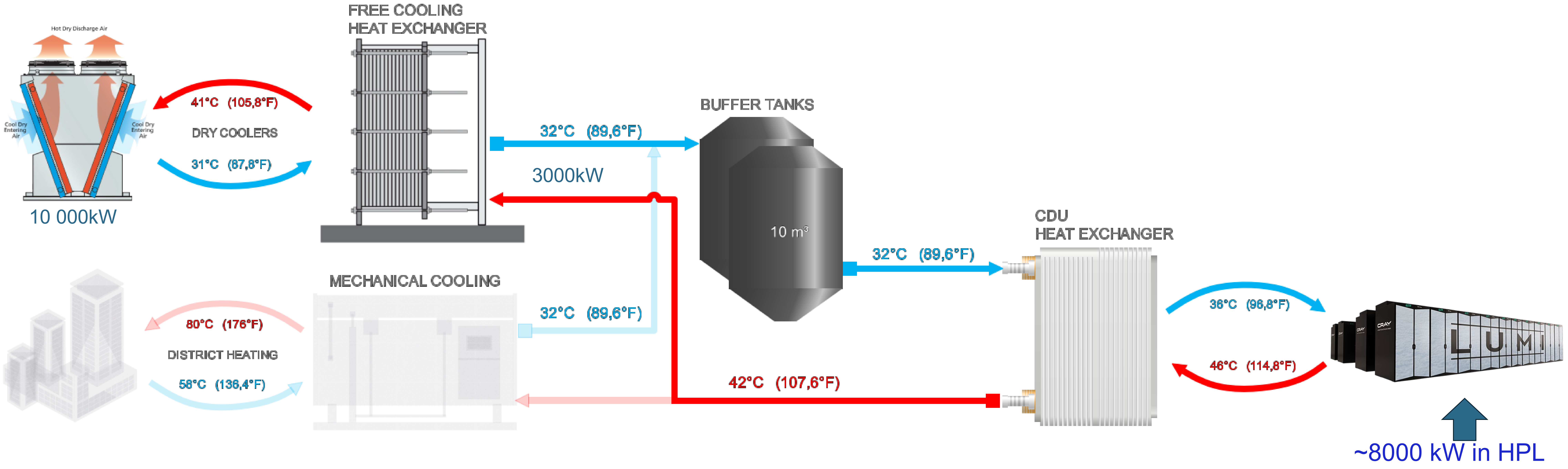
- **Partial free cooling with power limit**
  - Power limit is used if the district heating power plant indicates that there is not enough demand on the district heating network. Part of the heat load is bypassed to free cooling heat exchangers



LUMI Cooling and heat reuse

Three individually controlled cooling modules started in 2300kW steps based on LUMI IT power

- **Free cooling** with 3x 3300kW heat exchanger and total of 10 000kW DCR capacity
    - All three free cooling module heat exchangers are connected to the same main pipeline on rooftop DCRs
    - Free cooling is only operated when there is no demand on the district heating network or during a LUMI maintenance break



LUMI

- PUE (Power Usage Effectiveness) = Total Energy / IT Energy
- ERE (Energy Reuse Effectiveness) = Total Energy – Energy reuse / IT Energy
- ERF (Energy Reuse Factor) = Energy reuse / Total Energy
- COP (Coefficient Of Performance) = Energy reuse / Heat production Energy

- Electrical energy
- Thermal energy / heat

EUROHPC
SUMMIT 2024

ANTWERP
18–21 MARCH

TO EXASCALE
AND BEYOND

Julita Corbalan

Barcelona Supercomputing Center (BSC)

EuroHPC
Joint Undertaking

be
EU
belgium24.eu

ANTWERP

How does one define 'an Energy Efficient' HPC System?

From data to action

ANNOTATION

UNLEASHING THE POWER OF EUROPEAN HPC AND QUANTUM COMPUTING

Euro HPC monitoring strategies for energy efficiecy

Julita Corbalan (UPC/BSC/EAS)
EAR Project leader

DC power/energy in metrics

- Data Centers focuses on performance (GFlops)
- Energy efficiency (in Top500) is GFlops/watts
- Not all the jobs performance can be measured in GFlops
    - IO Jobs
    - Not computational intensive Jobs (or not vectorial instructions used extensively)
    - Others?

Energy efficiency

- Green 500 (Top 500) uses HPL, but Jobs do not report GFlops (or other metrics)
- To compare between DC we should compare with same workload
  - That's the idea of using HPL (is HPL enough for an energy efficiency classification ?)
- Is the energy efficiency of the DC
  - the energy efficiency of HPL?
  - Is the average of the efficiency of several pre-selected applications
  - is the average of 1 year of the DC workload execution?? (then not easy to compare, do we want to compare or just compute)

Energy efficiency

- Efficiency is always a ratio : performance vs power/energy consumption
- Performance is job performance
- Power/Energy could be per job or per cluster
- We need to collect job performance metrics automatically
- We need to collect cluster power metrics (power/energy) automatically
- And corelate them !!!!
  - Job GFlops/Watts (classic)
  - Workload GFlops/Watts
  - Throughput/Watts (generic)

What are we doing?
How much we consume to do it?

DCeP= Data Center EnergyProductivity
usefulworkproduced/ totalData Center energyConsumedproducingthiswork

Is it the same idea??

EAR customers strategy: BSC/MN5, SuperMUC-NG, Snellius

- At the job scope . Per job/Step/Node (also at runtime but maybe too much)
  - Job time and energy accounting for all the jobs.
  - Job performance metrics and detailed energy for all the jobs executed with EAR job manager (EAR library)
    - Time and Gflops but also CPI, Memory Bandwith, IO activity, MPI hints, GPU info (GPU GFlops is difficult to compute!!!)
- At the computational node scope: periodically (1-2 min), per node, even when idle
  - Power, CPU frequency, Memory frequency, temperature. Stored for some period and after that removed (2-3 months)
- Group of nodes/Cluster scope
  - Only power consumption
- No correlated with other system tools providing , for example, Storage information

Energy efficiency actions

- No pro-active actions taken in general
  - Users are not aware of their energy efficiency metrics! (even they are available and easy to collect)
  - No energy efficiency analysis and evaluation (audits) and the job level
  - No workload analysis at the DC scope
    - Who is consuming energy and to do what!
- The situation has started to change but (in my opinion) more motivated by some users/groups eco-responsability than from Data Center actions

Summary

- If we agree on energy efficiency = Work done/energy consumed to do it
  - We have to agree on what we consider as “work done”
  - Energy consumed it’s easy to agree on “including everything”
  - It must be easy to measure for DC owners
- If we want to improve it (reduce the energy to do the same work) we must report to users and make them aware of these metrics

EUROHPC
SUMMIT 2024

ANTWERP
18–21 MARCH

TO EXASCALE
AND BEYOND

EuroHPC
Joint Undertaking

be
EU
belgium24.eu

Bastian Koller

High Performance Computing Centre Stuttgart (HLRS)

EUROHPC SUMMIT 2024

ANTWERP

Energy Efficient Data Center Design and Operation @ HLRS

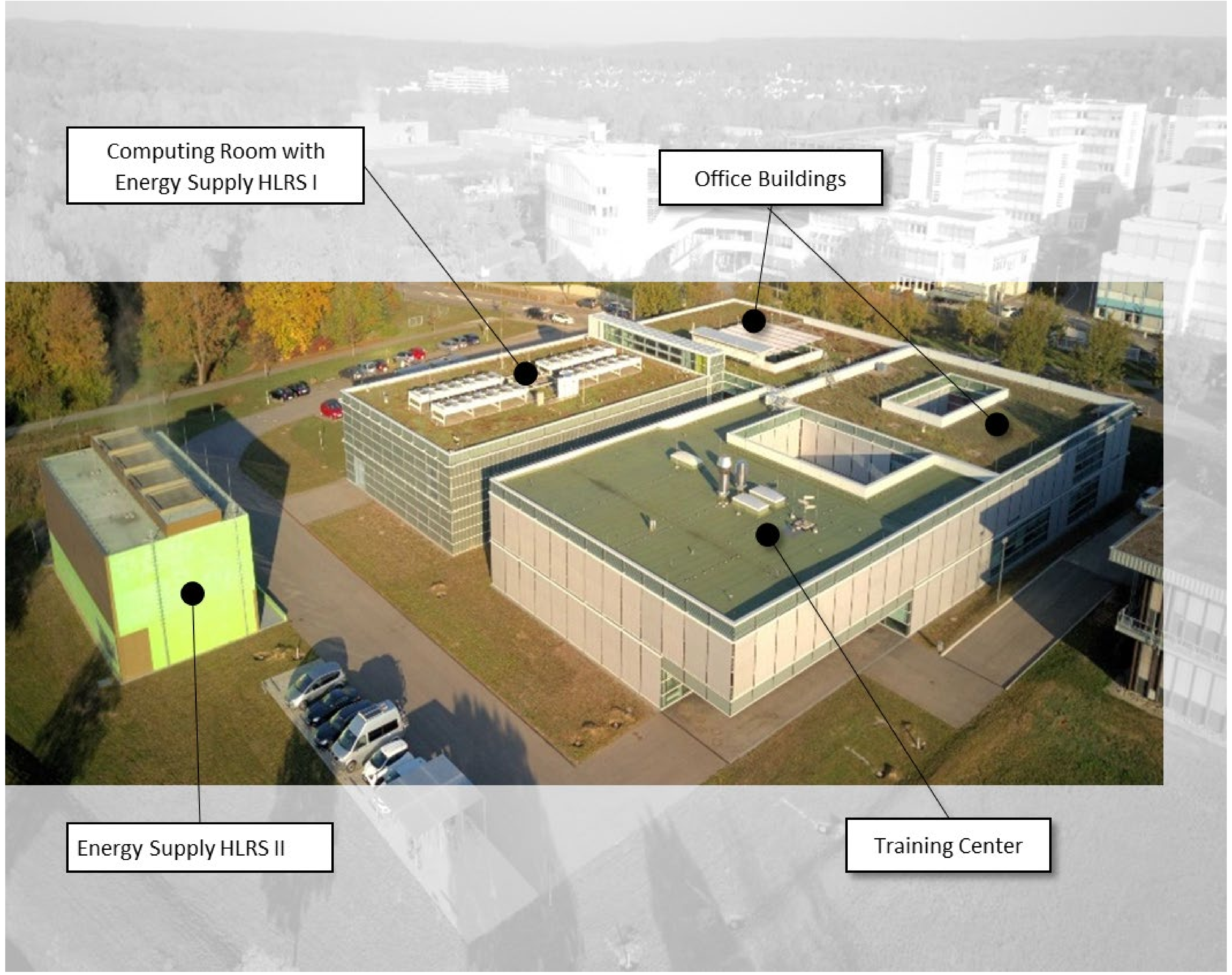
Bastian Koller, HLRS

UNLEASHING THE POWER OF EUROPEAN HPC AND QUANTUM COMPUTING

ANNOTATION

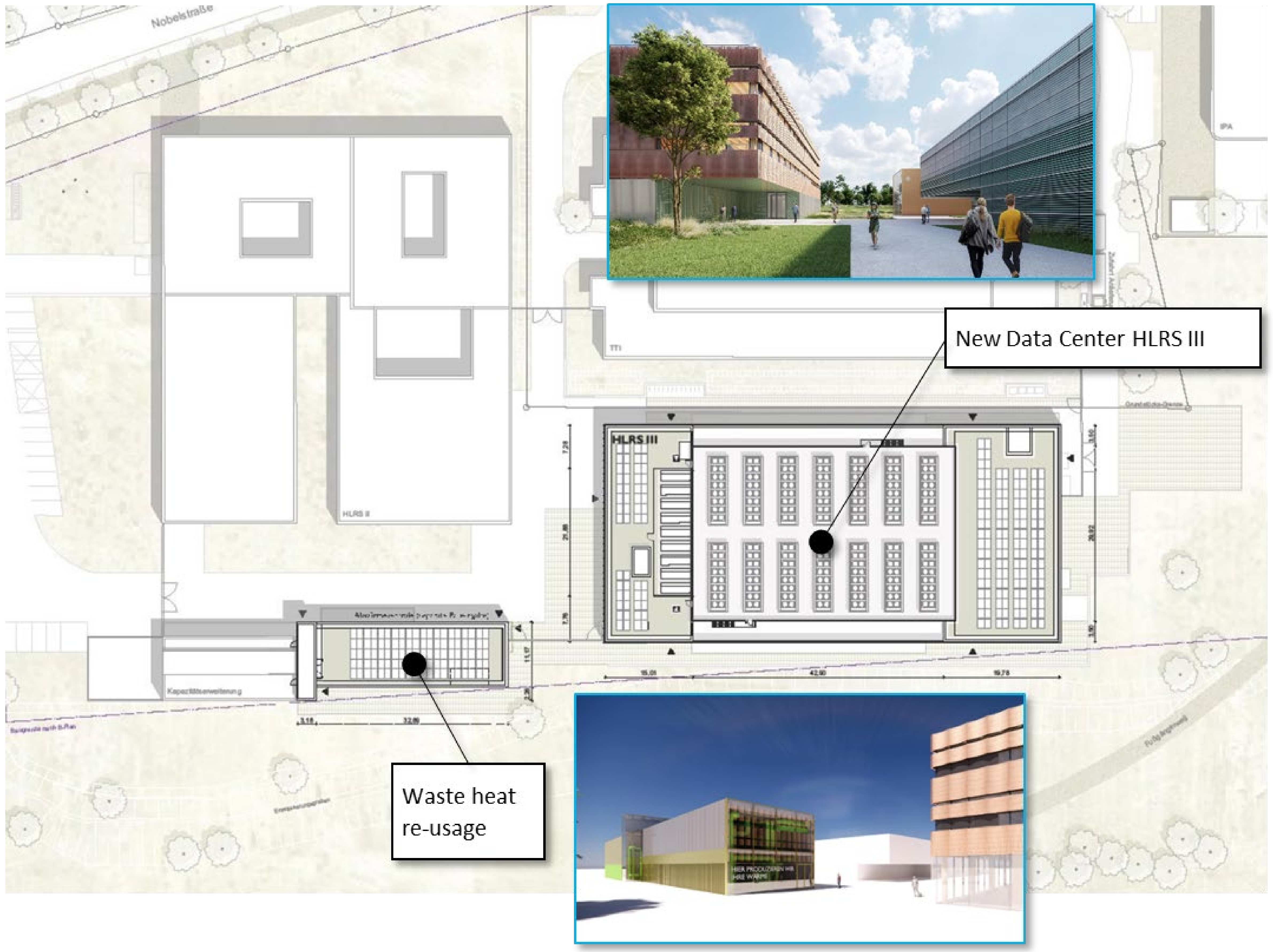
HLRS Infrastructure Today

- Energy Supplies
  - 4.6 MW maximum IT-Load
  - 860 kW battery supplied UPS
- Cooling Infrastructure
  - ~1 MW recirculating air cooling
    - 14 °C inlet temp. / 22°C return temp.
    - 50% free cooling (dry coolers)
    - 50% district cooling
  - ~4 MW water cooling
    - 25°C inlet temp. / 35°C return temp.
    - 85% free cooling (wet cooling towers)
    - 15% district cooling



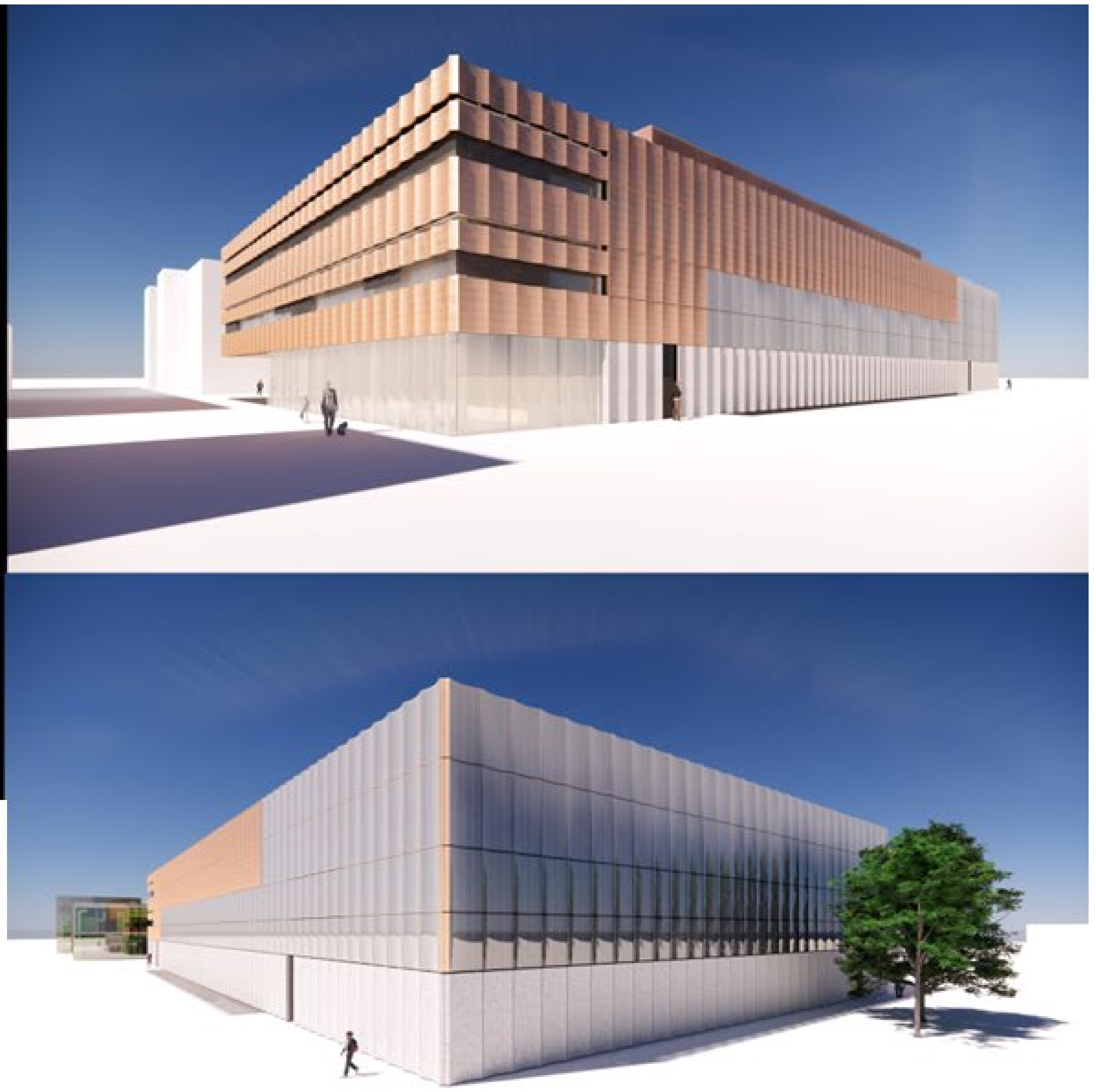
New Data Center HLRS III - Parameters

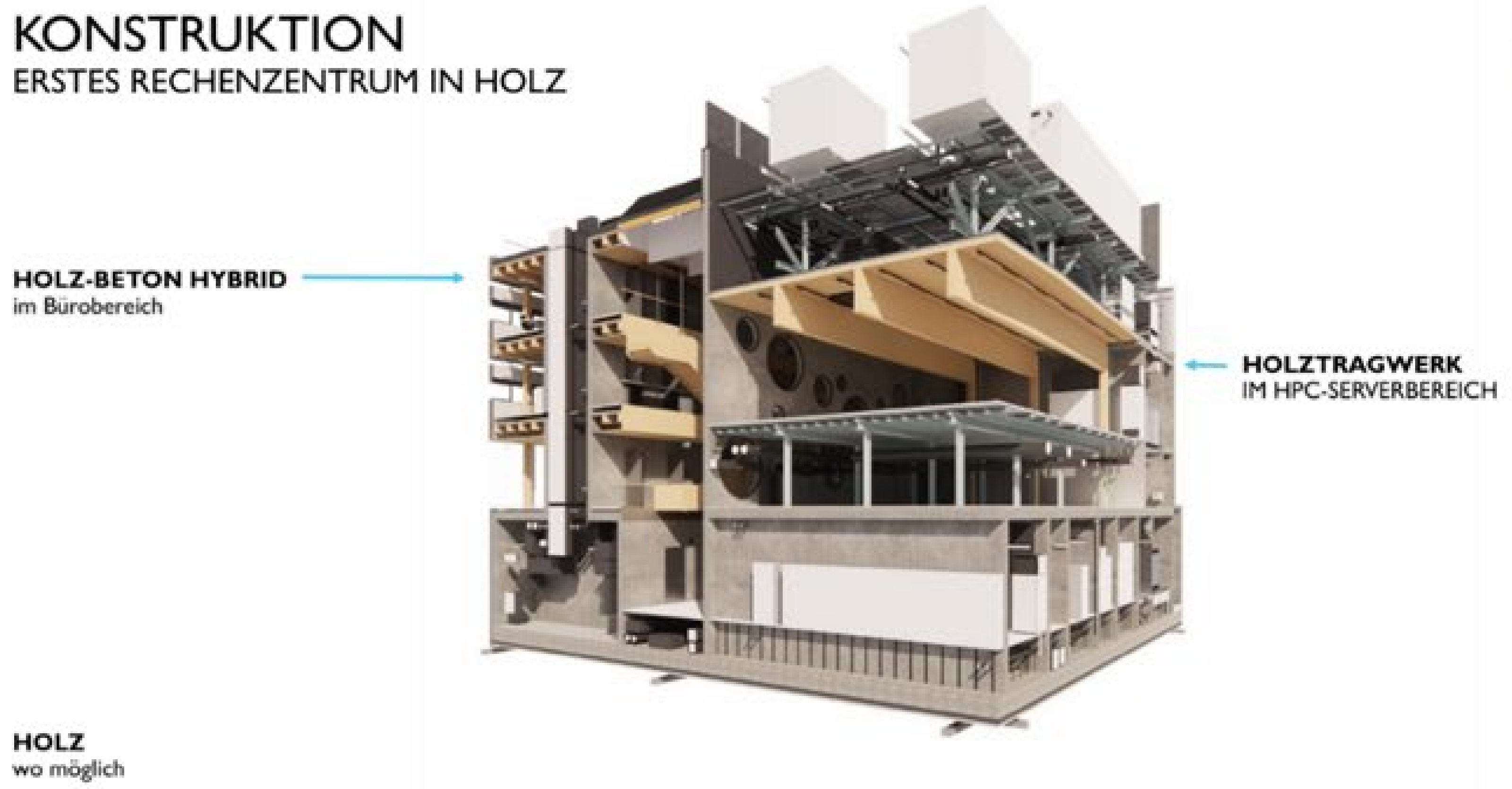
- Energy Supplies
  - 8.0 MW maximum IT-Load
  - 0.6 MW battery supplied UPS
  - 7.4 MW Fly Wheel UPS
- Cooling Infrastructure
  - 0.8 MW recirculating air cooling
    - 18 °C inlet temp. / 24°C return temp.
  - 7.2 MW water cooling
    - min. 32°C inlet temp.
    - min. 42°C return temp.
  - No district cooling
  - 100% free cooling (adiabatic dry coolers)



New Data Center HLRS III – Energy Efficiency

- Energy savings and sustainable concepts are key features
- Photovoltaic Façade & Roof
  - ~ 850 m²
  - 100% self-consumption
  - Energy output to be estimated
- Wooden Construction where possible
  - Computing room ceiling
  - Office Section





New Data Center HLRS III – Waste Heat Re-Usage

- Waste heat re-usage is one key component for the decarbonization of the Campus Vaihingen University of Stuttgart
- Innovative Cooling Concept
- No external cold water loop
- All heat is emitted with at least 42°C (target is 50°C)
- 6.2 MW average waste heat re-usage
- Estimations for the Campus Vaihingen:
  - 20 % reduction of total energy consumption
  - 50% reduction of CO_2 emissions

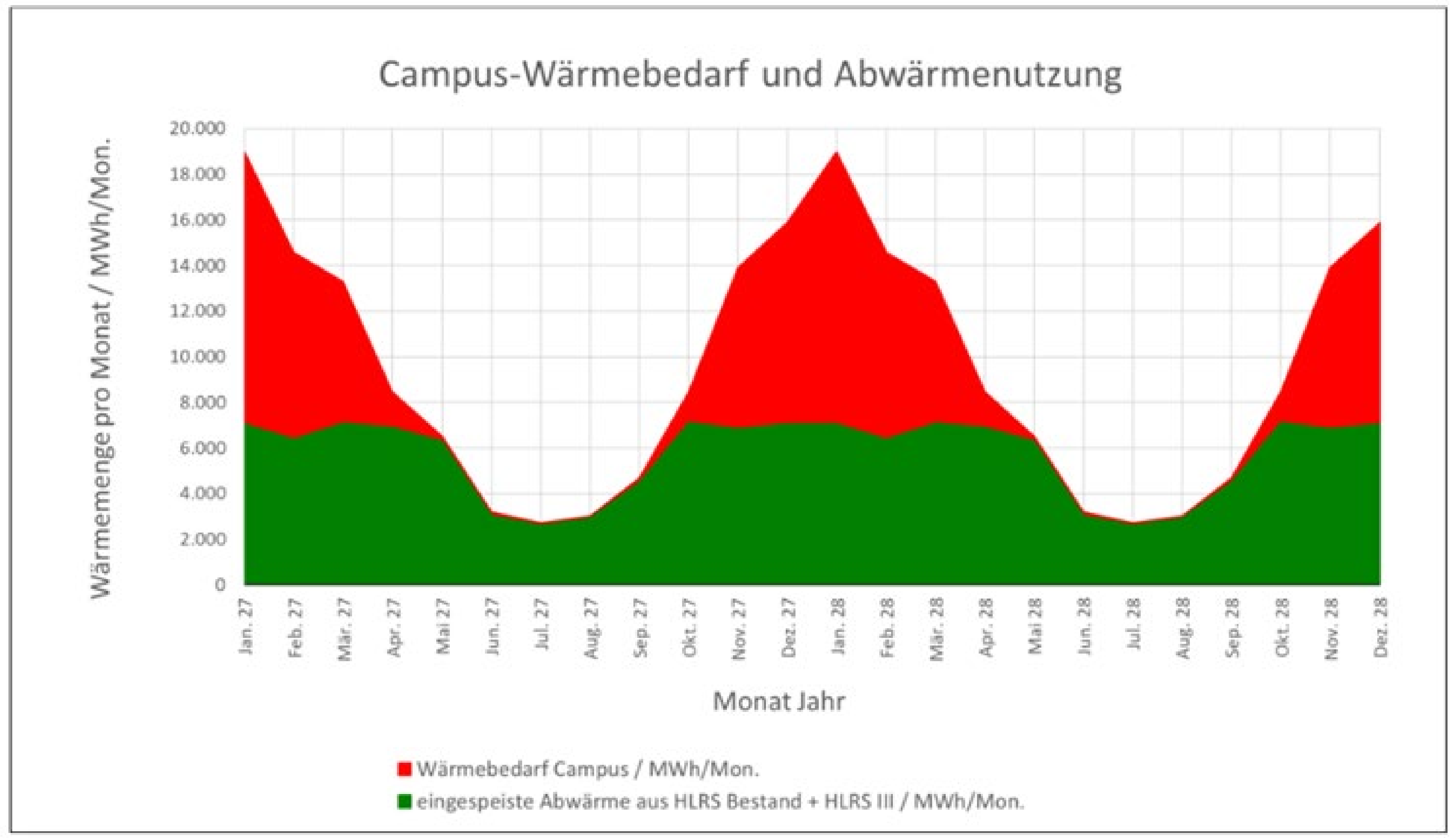
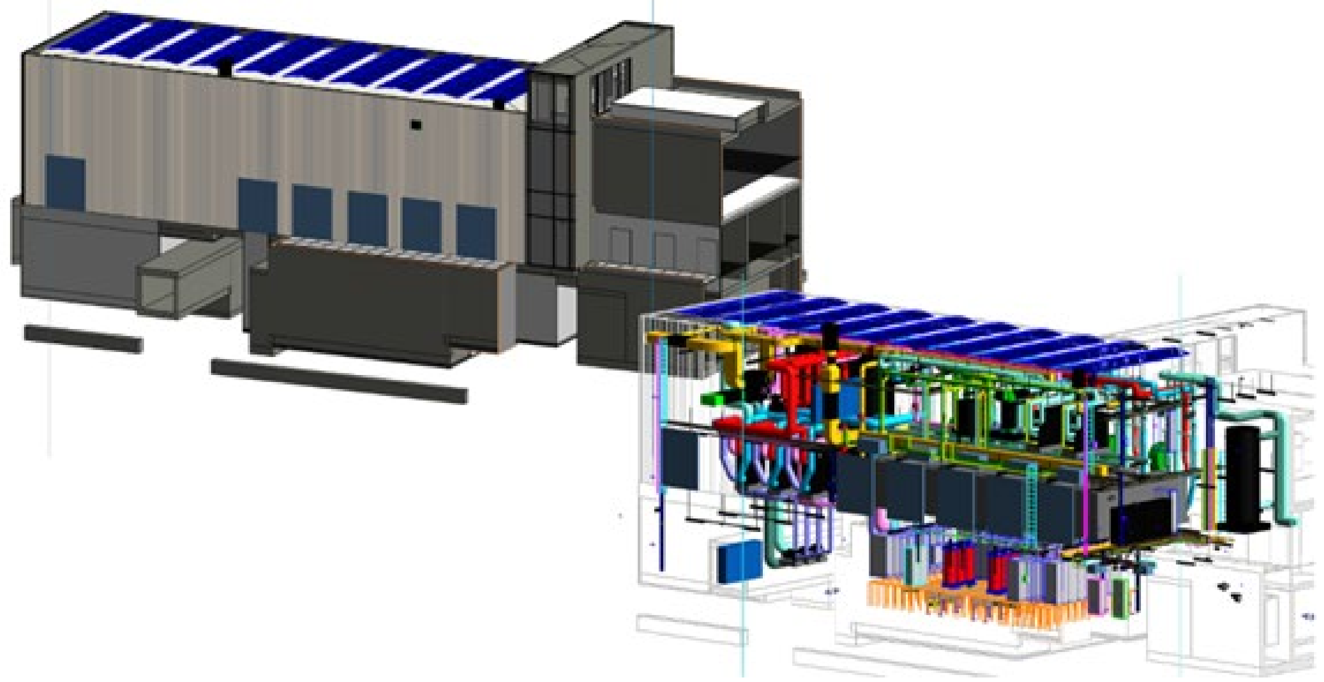


Abbildung 2: Prognostizierte Abwärmenutzung innerhalb Fernwärmenetz



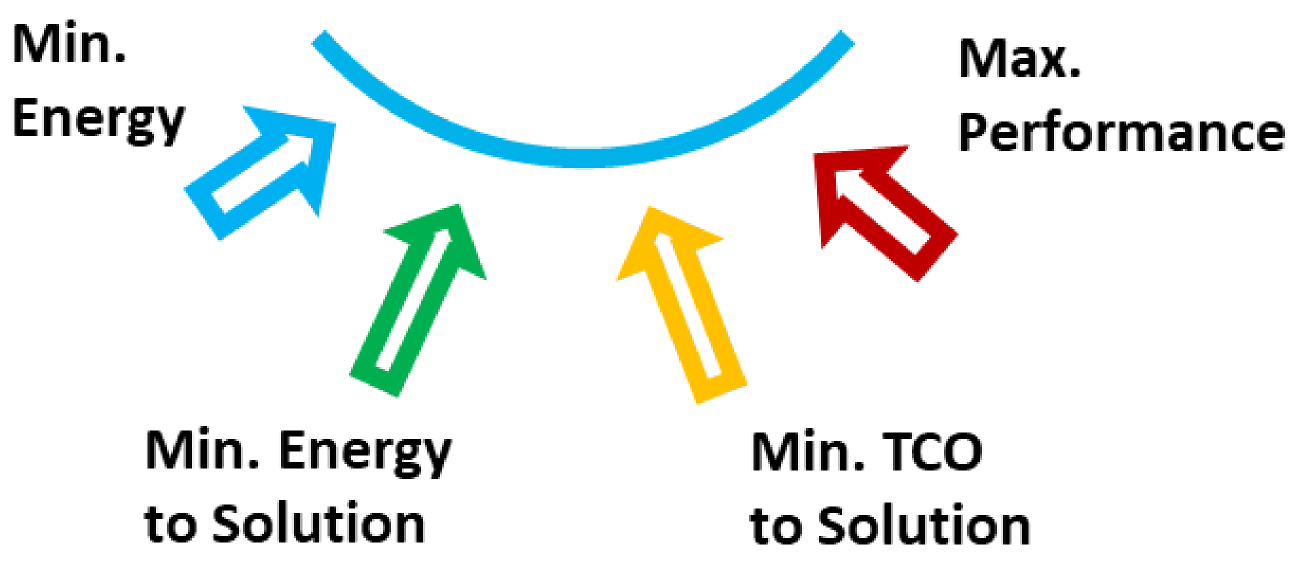
Energy Efficient Operation in Production

- High energy prices
- Over-provisioned Systems
- Static power caps

Energy Optimization
- Use system metrics only
- Aim for minimal overhead
- Assign optimal power level
- React to different application phases
- Transparent
- w/o user input.

Hewlett Packard Enterprise

PowerSched



Dynamic Power Steering depending on
- system load
- external power or cooling constraints

Thank you!

EUROHPC
SUMMIT 2024

ANTWERP
18–21 MARCH

TO EXASCALE
AND BEYOND

Mentimeter

Next Round

Join at menti.com | use code 3876 9437

Mentimeter

Final Round

Join at menti.com | use code 3876 9437

Take-away Message

It is possible to improve!

Start monitoring now!!

But don’t leave it at that!!!

- Analyse the data, share it, draw your conclusions and act on them
- EuroHPC could really help here!
- Get ready for this, you might have to do this soon
- (Standardisation would be very helpful)