Scaling Sustainable Computing

New Methods for Enhancing Efficiency in the SKA Regional Data Centers

Dr. Denisa Constantinescu, Darong Huang, Prof. David Atienza

Embedded Systems Laboratory (ESL) EPFL, Switzerland

denisa.constantinescu@epfl.ch

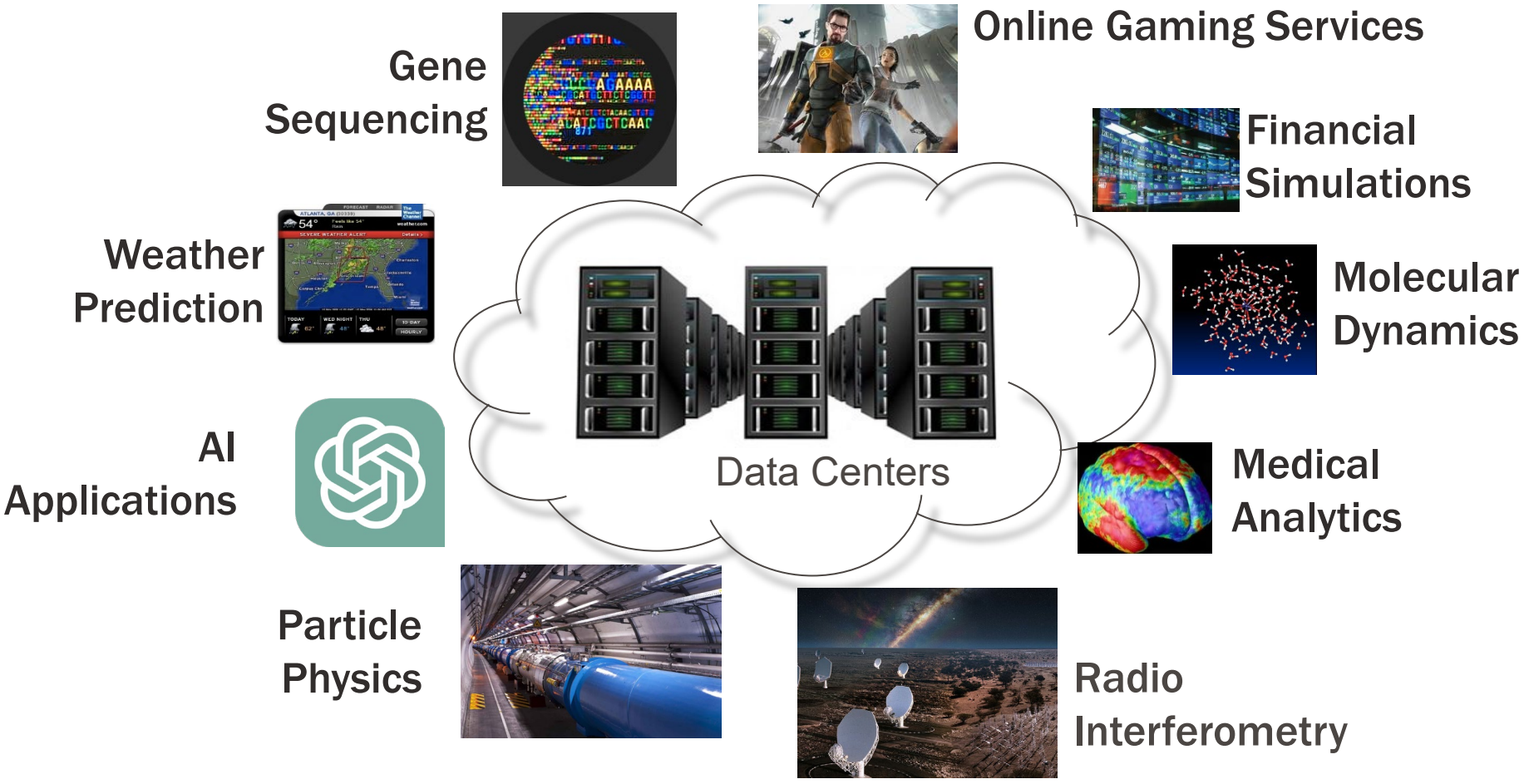
Computing is Ever More Indispensable

Key in supply-chain of products, services and science

...and Growing!

but not sustainability-driven

cannot scale without improving efficiency



Shift towards **sustainable computing**

- [x] 14th of March ‘24
  [EU Data Centres Reporting Scheme]

→ **KPIs to rate sustainability**

- **Total energy, water, heat in DCs**
- **Wasted, reused, renewable sources**

- [ ] EU’s target [Fit for 55]

→ **reduce net GHG** emissions by at least **55% by 2030**

GREEN500 LIST - JUNE 2024

R_{max} and R_{peak} values are in PFlop/s. For more details about other fields, check the TOP500 description.

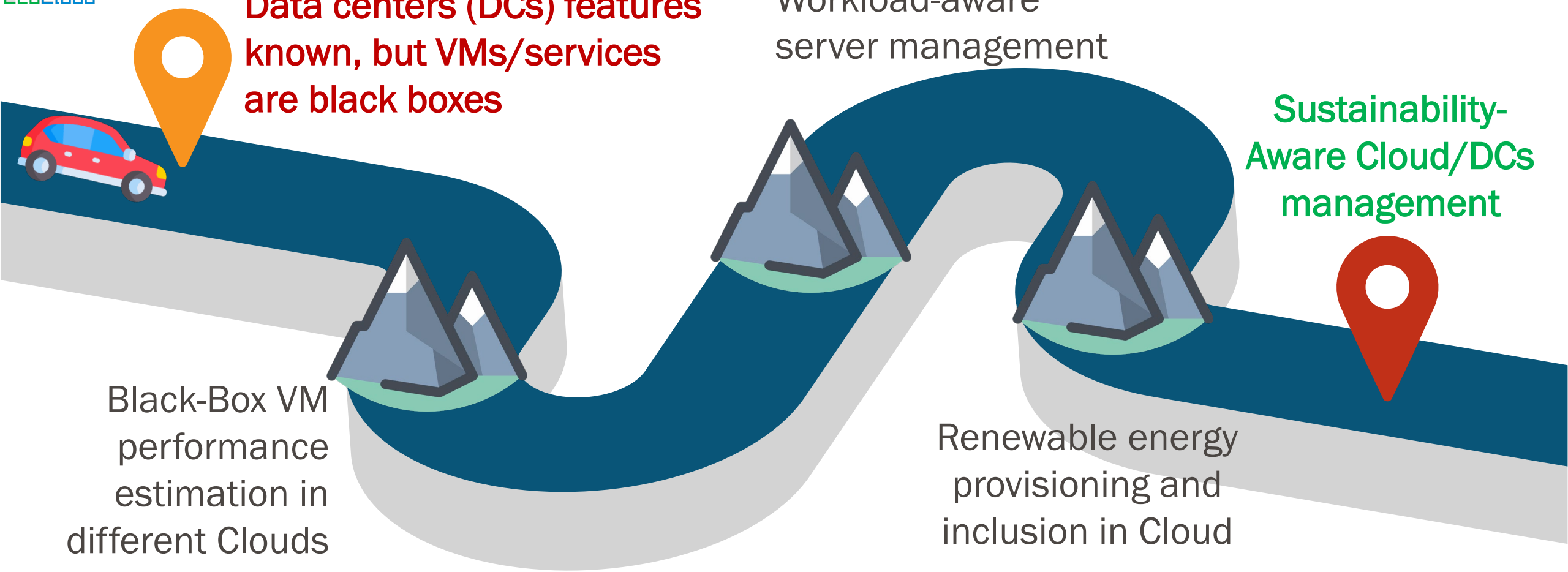
R_{peak} values are calculated using the advertised clock rate of the CPU. For the efficiency of the systems you should take into account the Turbo CPU clock rate where it applies.

Green500 Data

← 1-100 | 101-200 | 201-300 | 301-400 | 401-500 →

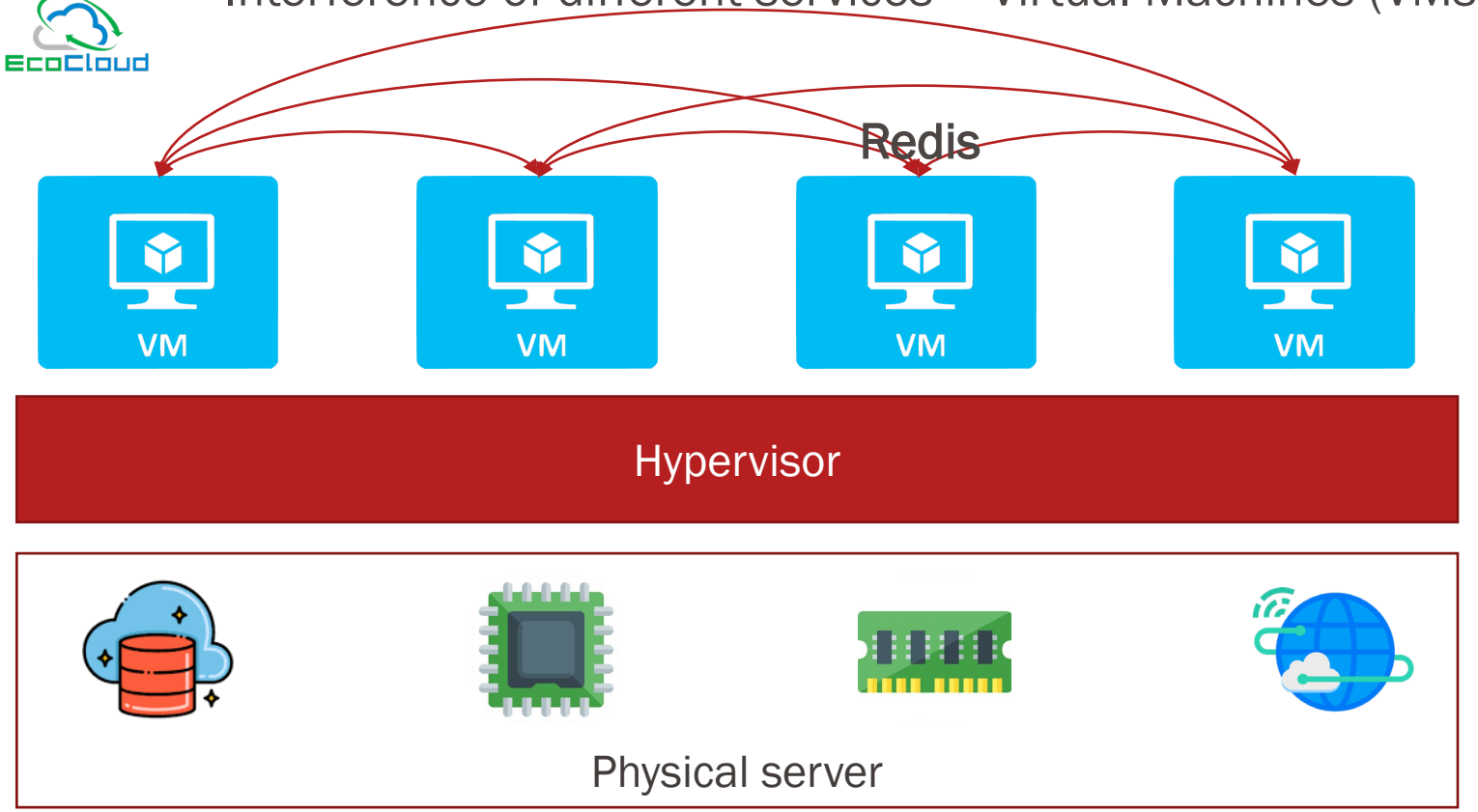
Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)
1	189	**JEDI** - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, ParTec/EVIDEN EuroHPC/FZJ Germany	19,584	4.50	67	72.733
2	128	**Isambard-AI phase 1** - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE University of Bristol United Kingdom	34,272	7.42	117	68.835
3	55	**Helios GPU** - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE	89,760	19.14	317	66.948

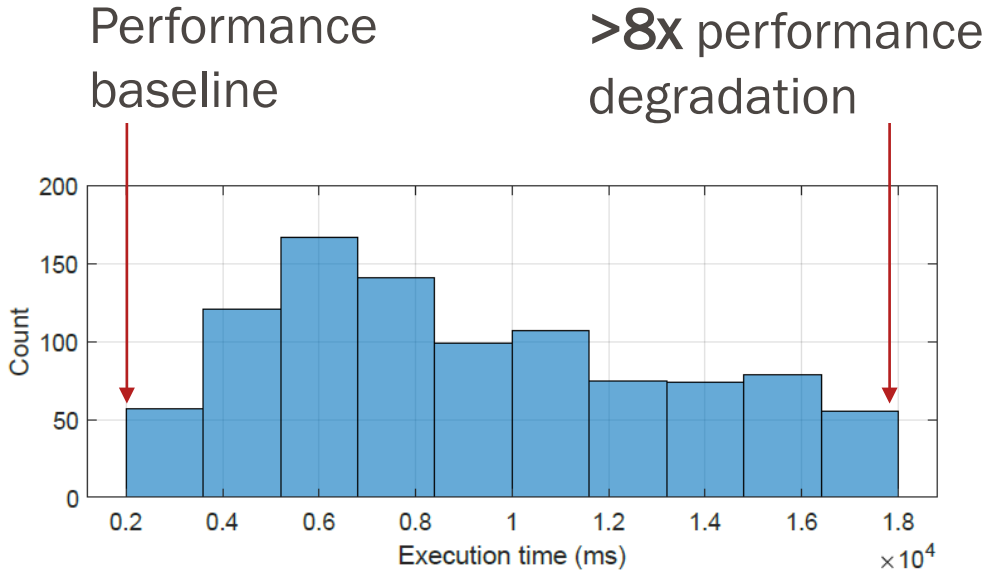
Challenges in our Path towards Sustainable Cloud/DCs



Interference Problem on (Virtualized) Cloud Services

Interference of different services – Virtual Machines (VMs)

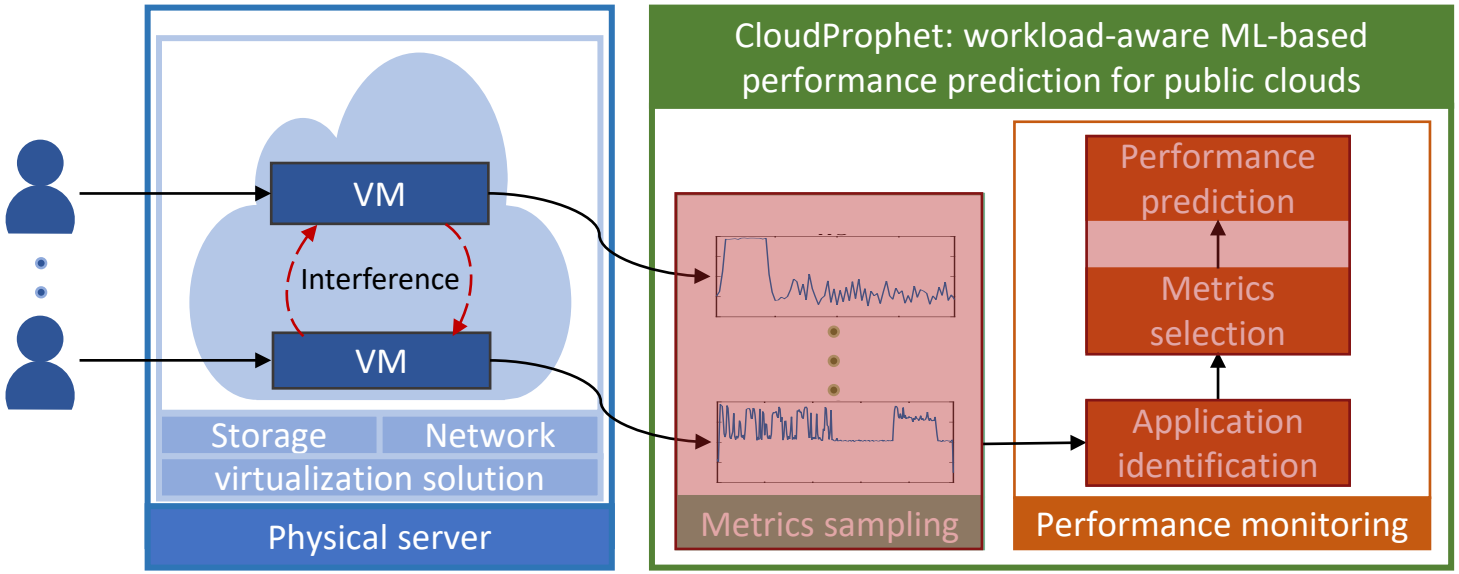




Performance of Redis benchmark

Collocated black-box VMs can suffer from severe performance degradation!
(Or over-provisioning needed to "guarantee" performance)

CloudProphet: Black-Box VM Performance Management



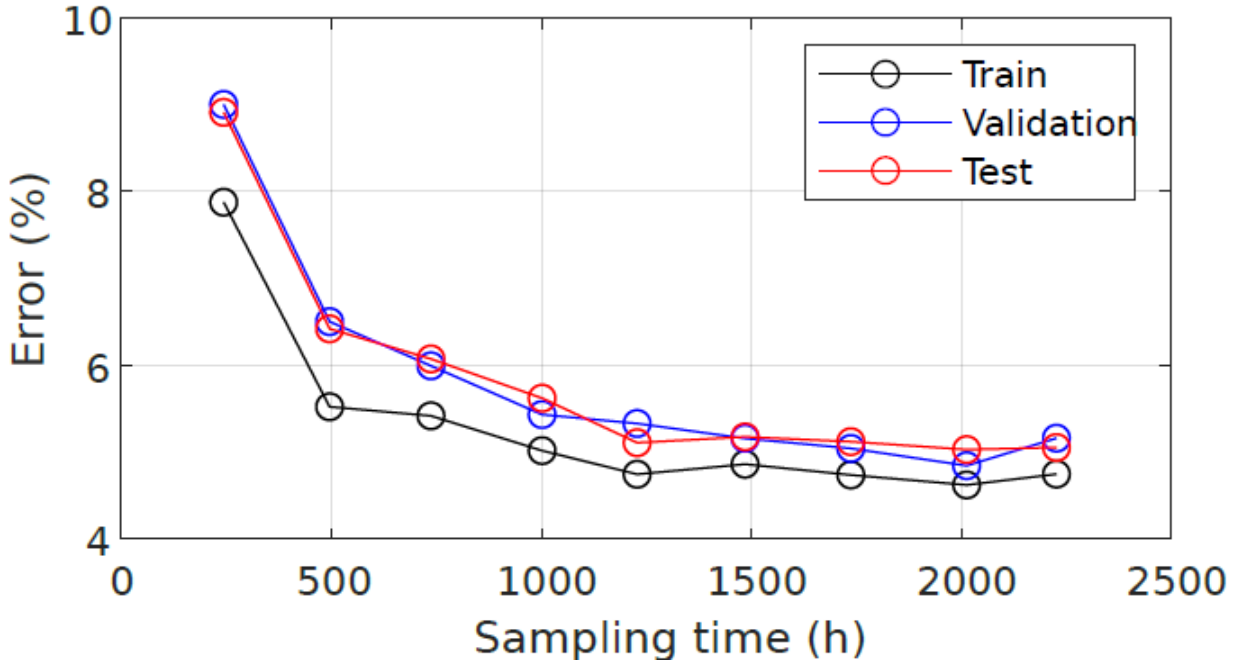
CloudProphet on IEEE
[Huang et al., TSUSC 2024]

CloudProphet-Dataset repo

Main steps:

1. Monitoring data (black box)
2. Application identification
3. Performance prediction

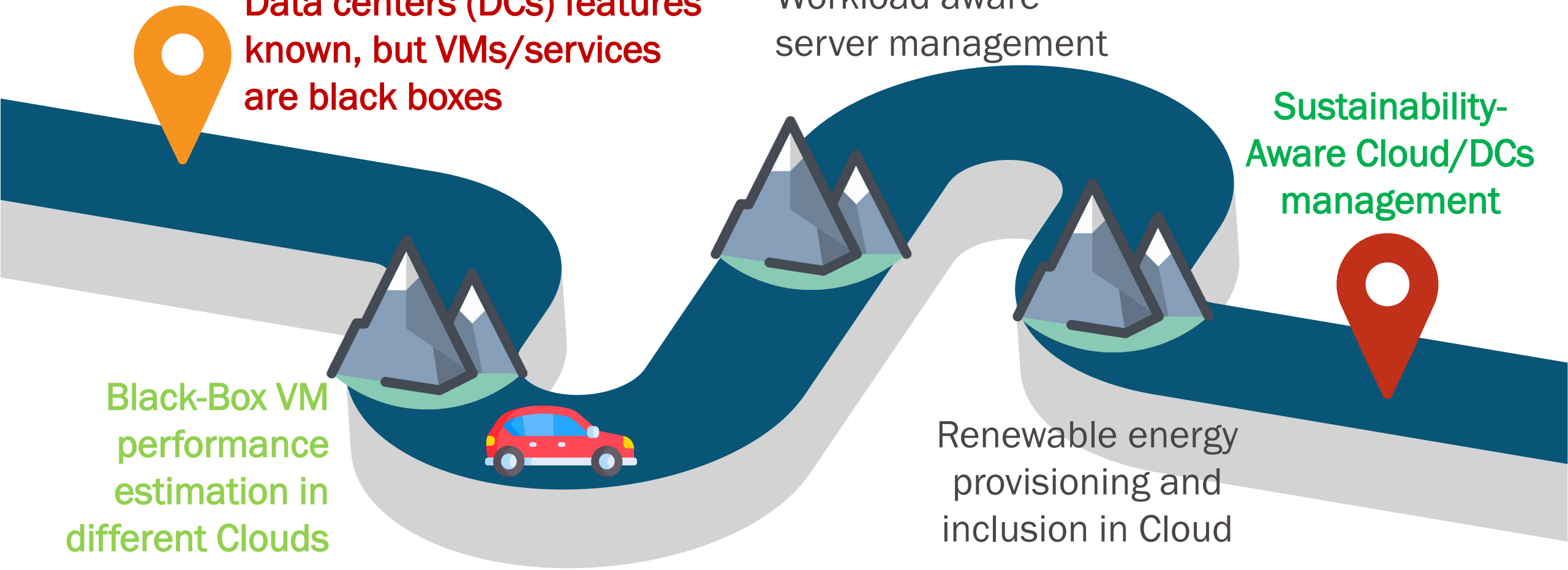
Accurate Performance Prediction of CloudProphet



Trade-off between sampling time and prediction accuracy

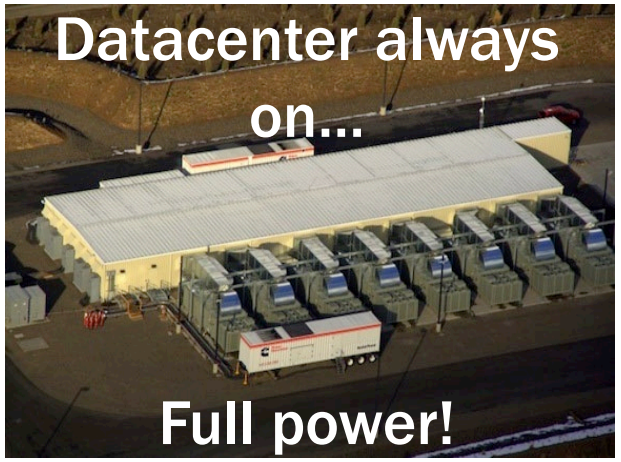
- **Less than 7% prediction error** after 20 days, better with more samples
- **5% performance prediction error** after 2 months of operation

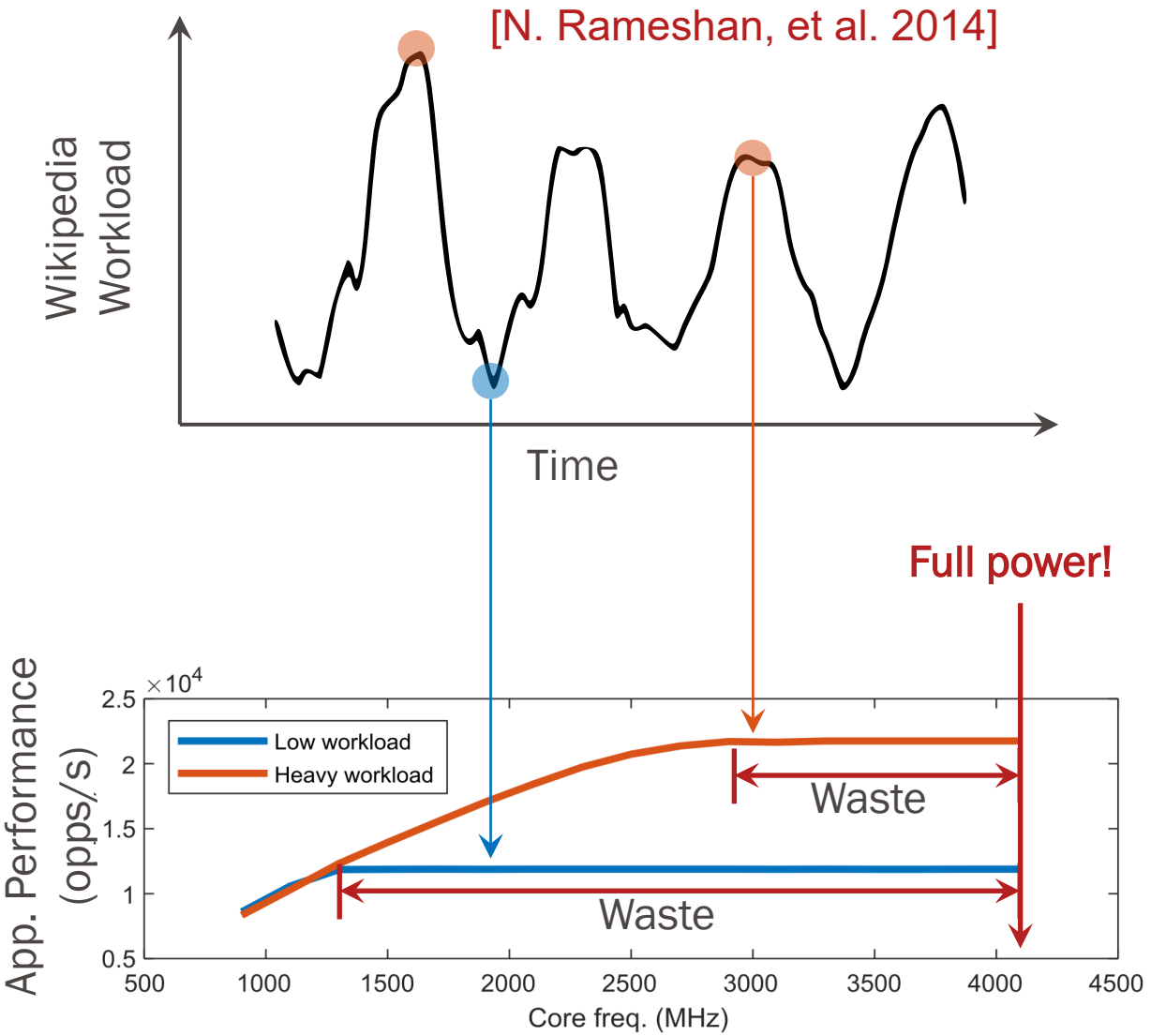
Challenges in our Path to a Sustainable Cloud



Large Resource Wasted in Cloud Designs!

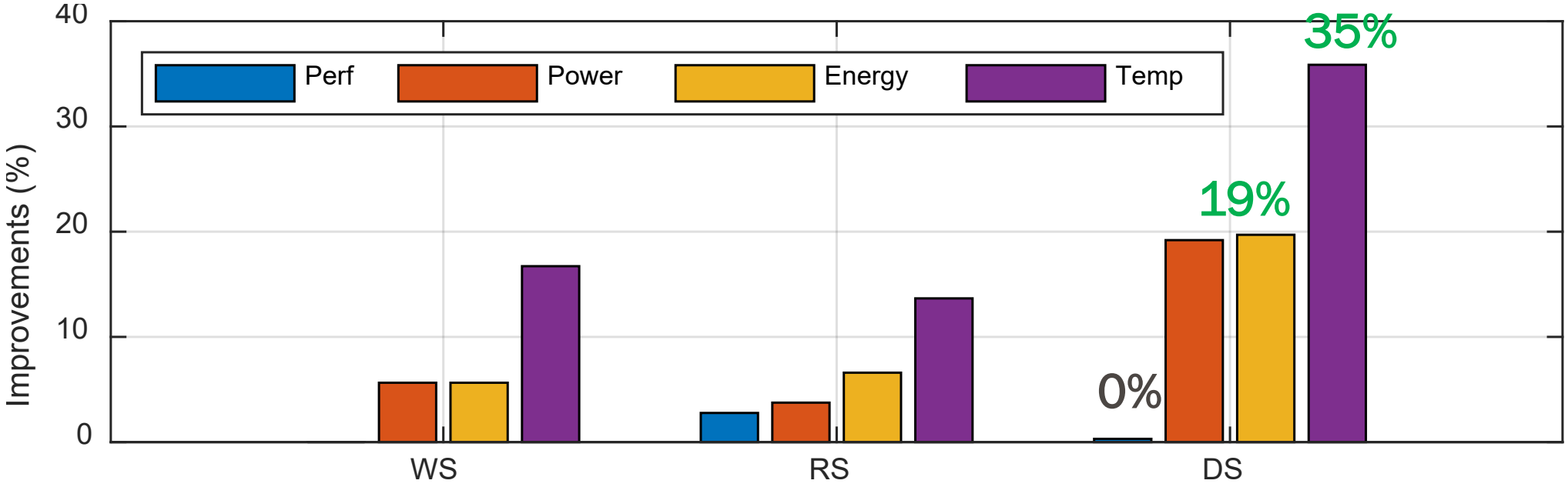
- Worse case resource provisioning paradigm: variable demand





- Hint: Appropriate **frequency scaling approach (DVFS)** can significantly reduced wasted energy in data centers
- **Scaling governors are critical!**

Improved Energy/Temp.-Efficient Server Management

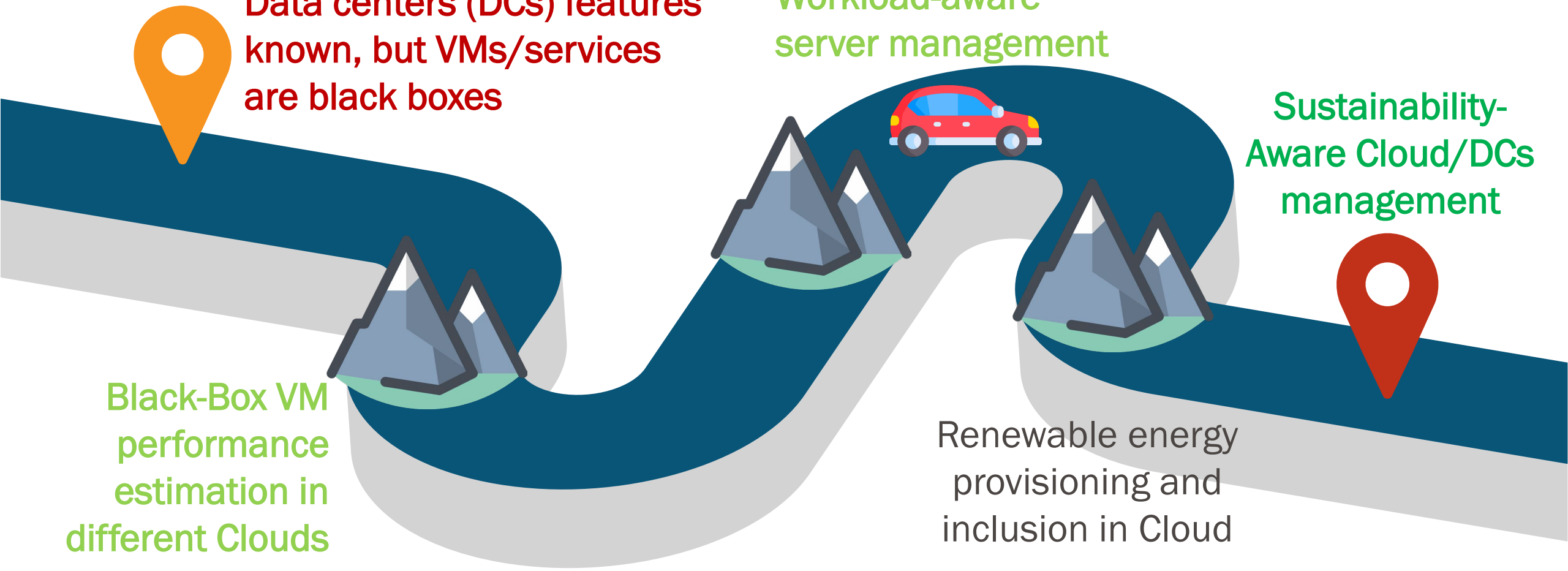


Compared with *intel-governor*, GreenDVFS achieves:

- No performance loss
- Up to **19% less energy** consumed
- Up to **35% lower temperature** in operation

GreenDVFS (Camera-ready)
[Huang et al., CCGrid'24]

Challenges in our Path to a Sustainable Cloud



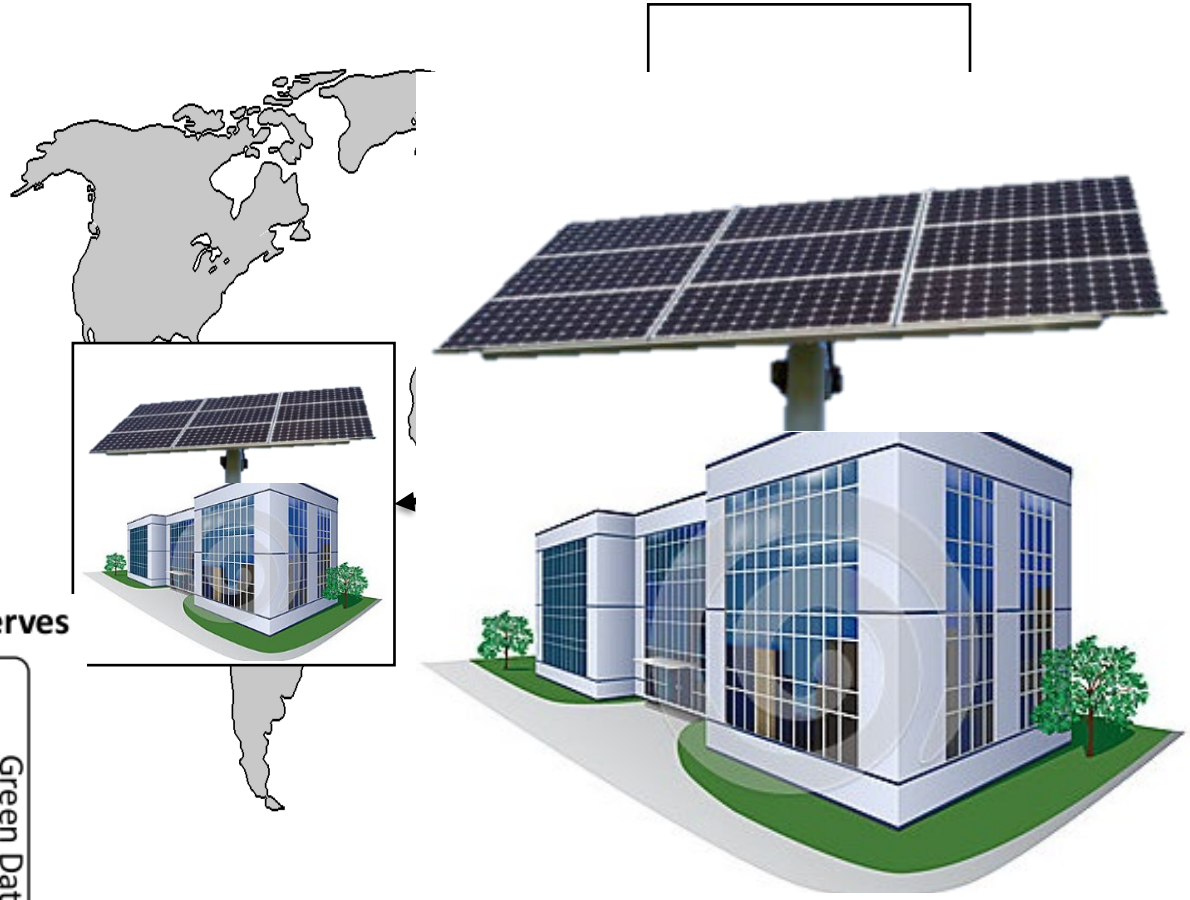
SRCNet: Sustainable Energy Sources in DCs Location

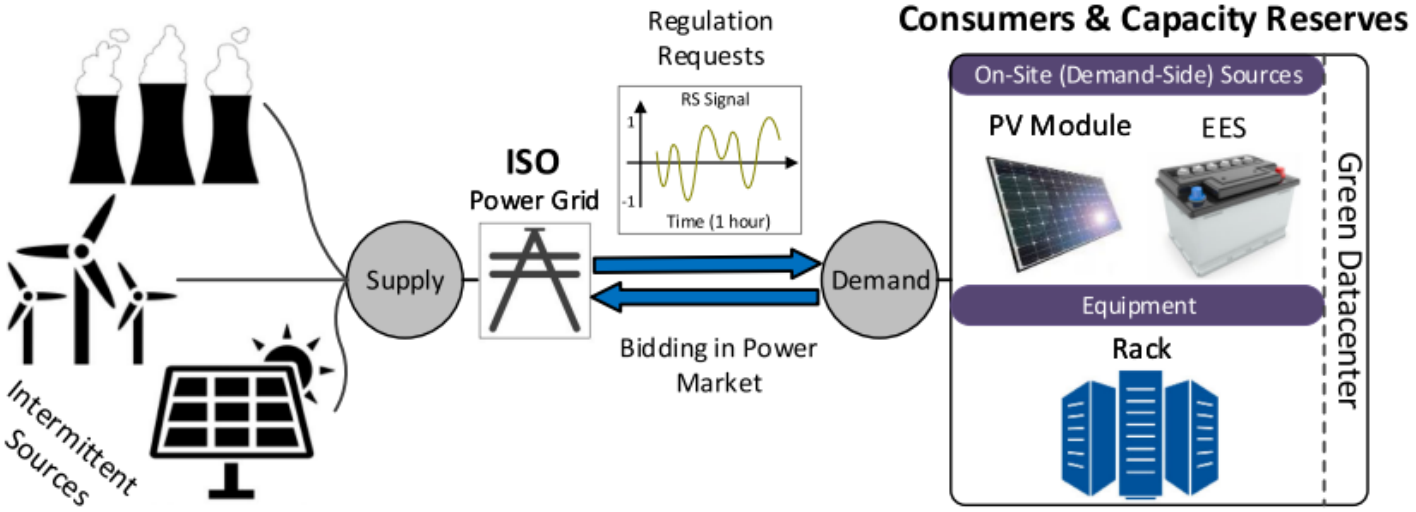
Geo-distributed data centers (DCs)

- Multiple DCs in different locations connected through network
- How to allocate VMs to different DCs?

Ideal placement for green DCs

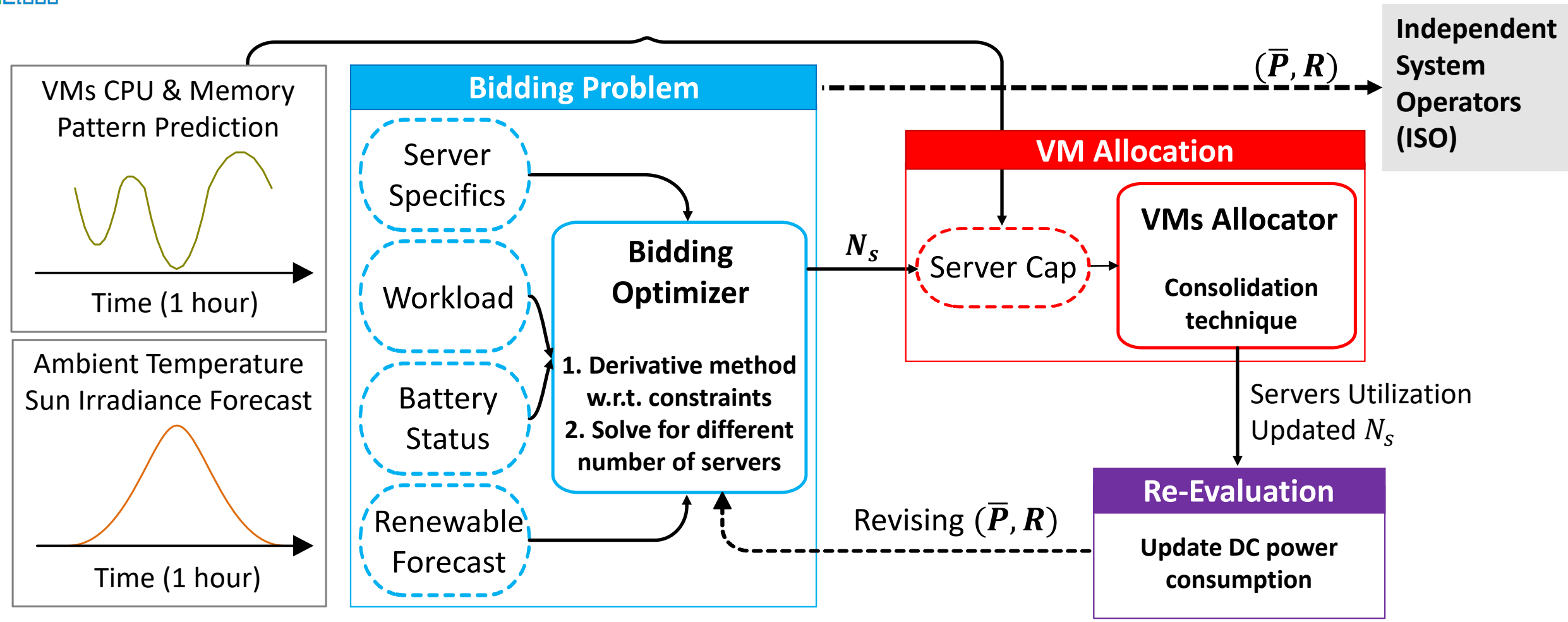
- How to manage renewable energy sources?



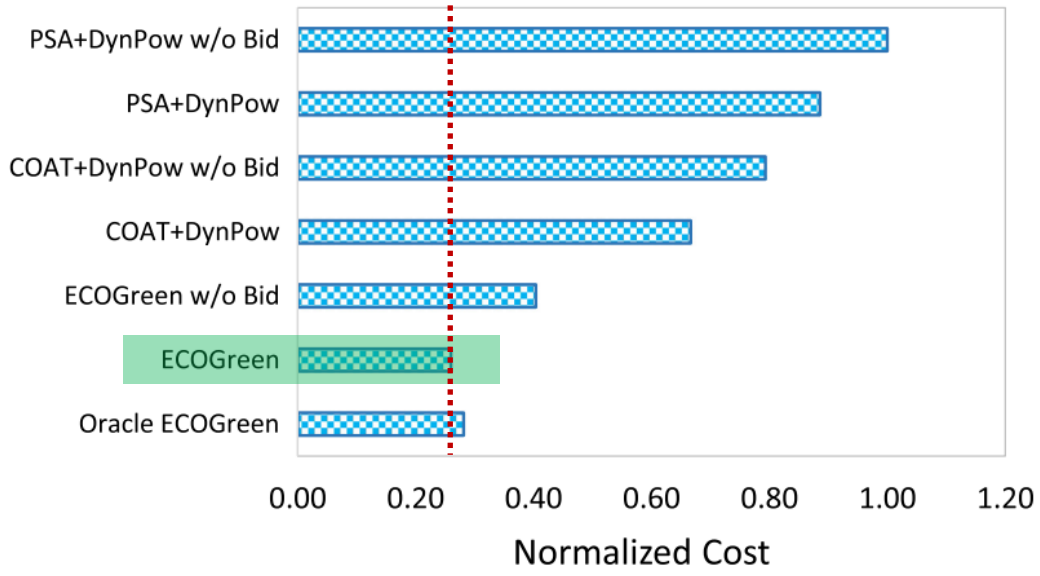


ECOGreen: Proposed Strategy

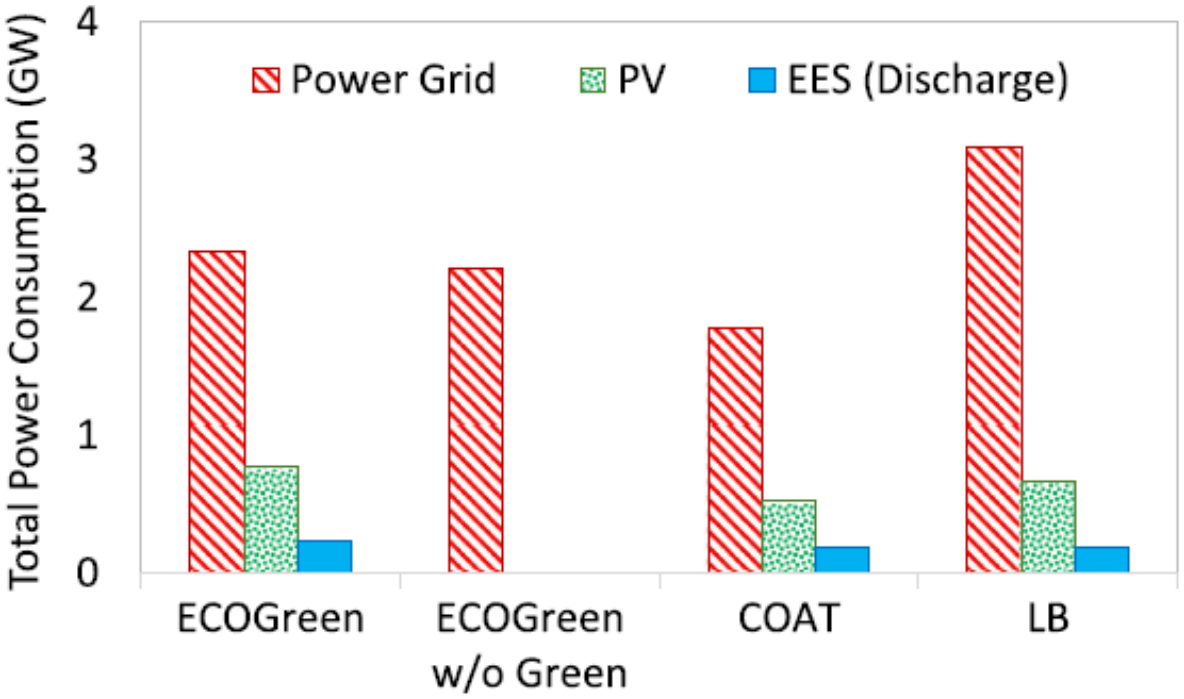
Hour-ahead power market (bidding)



ECOGreen: results for Green and Sustainable DC



Normalized monetary cost (1-week time horizon)



Different power supply sources (1-week time horizon)

- In comparison to the-state-of-the-arts, ECOGreen
  - **71% reduction** of financial costs
  - **48% increase of use in** renewable energy (more sustainable!)

Smmary: Sustainability-Aware Renewable Energy Management

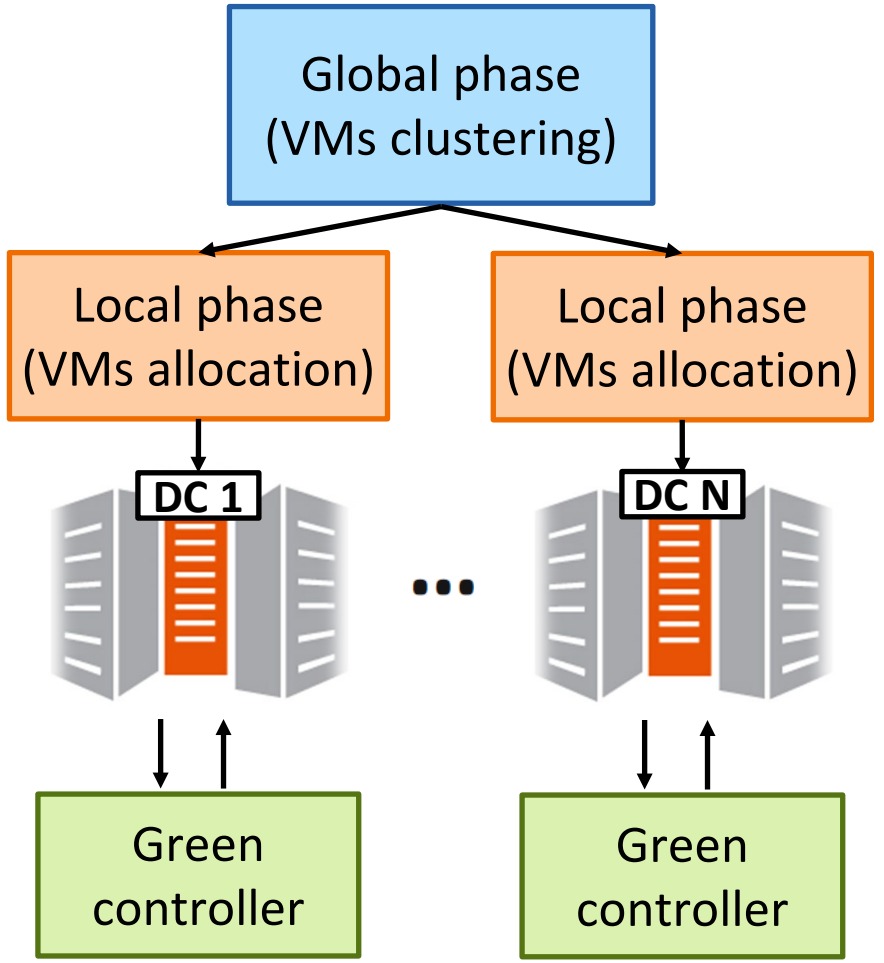
DCs/VM manag. (CloudProphet + GreenDVFS)

- **Global phase**: clustering VMs into DCs
- **Local phase**: VMs allocation for minimum performance degradation

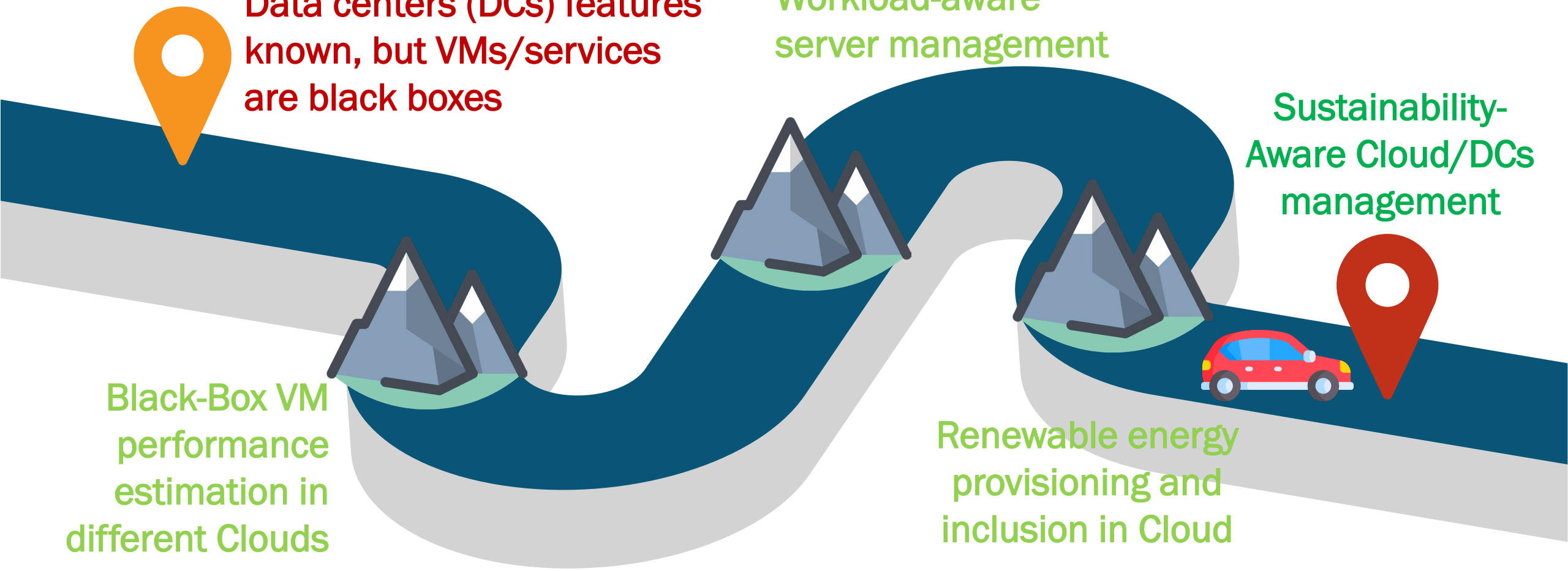
ECOGreen: Low-complexity green energy controller

- Management of renewable energy
- Add batteries in DCs: charge / discharge decisions

ECOGreen Energy Controller
[Pahlevan et al., TSUSC 2020]



Challenges in our Path to a Sustainable Cloud



Scaling Sustainable Computing in Networked DCs: **SRCNet**

IT/Cloud has enabled our progress for 50+ years
…but current **cloud systems are not sustainable**

How can we measure/estimate?

- Perf. upcoming workloads → CloudProphet
- Energy in DCs → work in progress
- Water & heat waste → work in progress
- GHG emissions → work in progress

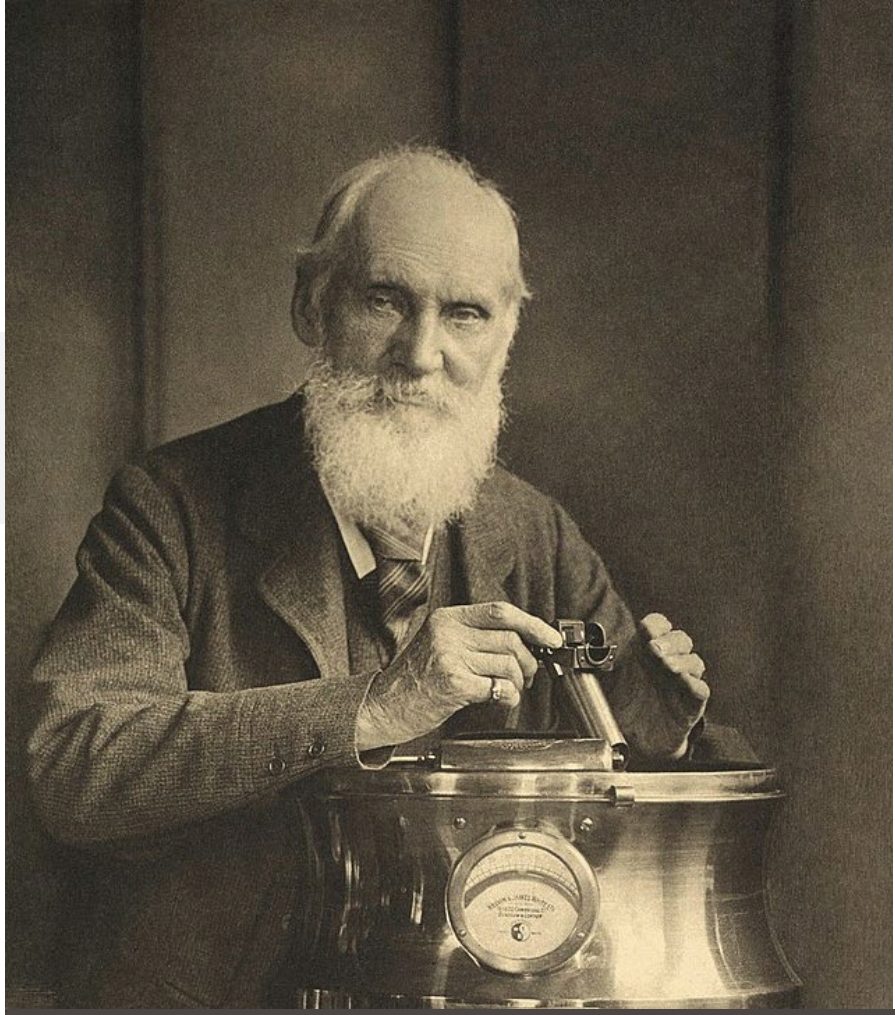
How can we improve sustainability?

Server level

- Power management → GreenDVS
- Specialized computing architectures → SEAMS

(multi)Datacenter level → EcoGreen

- Energy mix → very limited use of renewable
- Climate change impact → work in progress



"If you cannot measure it, you cannot improve it"
– Lord Kelvin, 1824 - 1907

Follow-up Proposal

Explore Sustainable Computing Techniques for SRCNet at EcoCloud

EcoCloud Sustainability Center ecocloud.epfl.ch

- ~150 m² of space for experiments on sustainable computing
- 50KW per rack/2.5m rack
- Monitoring: energy, temp., vibrat., humidity
- Cooling: air or water cooling (single-/2-phase)

Supports large multi-country academic-industry research projects on cloud

- **UrbanTwin:** An urban digital twin for climate action for Lausanne
- **Heating Bits:** DCs integrating heating and cooling supply of local districts
- **SEAMS:** Sustainable and energy-aware methods for SKA → seams-project.com

Microsoft HUAWEI intel Hewlett Packard Enterprise IBM Research Europe DeepSquare

ORACLE Meta AXA IMO Ventures infomaniak corintis JJ COOLING INNOVATION GEMATEG

Scaling Sustainable Computing

New Methods for Enhancing Efficiency in the SKA Regional Data Centers

Dr. Denisa Constantinescu

Embedded Systems Laboratory (ESL) EPFL, Switzerland

denisa.constantinescu@epfl.ch

Acknowledgements (alph. order): Prof. David Atienza, Prof. Ayse K. Coskun, Dr. Luis Costero, Mr. Darong Huang, Dr. Ali Pahlevan, Mr. Amirhossein Shahbazinia, Prof. Marina Zapater

Swiss National Science Foundation: SEAMS

seams-project.com

ETH Board: UrbanTwin JI Action

European Commission

European Research Council

Backup slides

Summary Tools

- IT/Cloud has enabled our progress for 50+ years
  - Multi-core servers and data centers are becoming more powerful
  - Big Data + Exascale era could be conceived…
- But current cloud systems are **not sustainable**; changes needed
  - Very different and dynamic workloads than classical HPC
  - Severe performance interference among VMs collocated together
  - Very limited use of renewable energy supplies
- ML-based management of DCs **to start conceiving a sustainable cloud**
  1. **CloudProphet:** Accurate and adaptive to new workloads (<7% accuracy error)
  2. **GreenDVFS:** Higher energy efficiency per server (20% less energy, 35% less temp.)
  3. **ECOGreen:** Multi-DC management + renewables (48% increase of renewables)

CloudProphet

GreenDVFS

ECOGreen

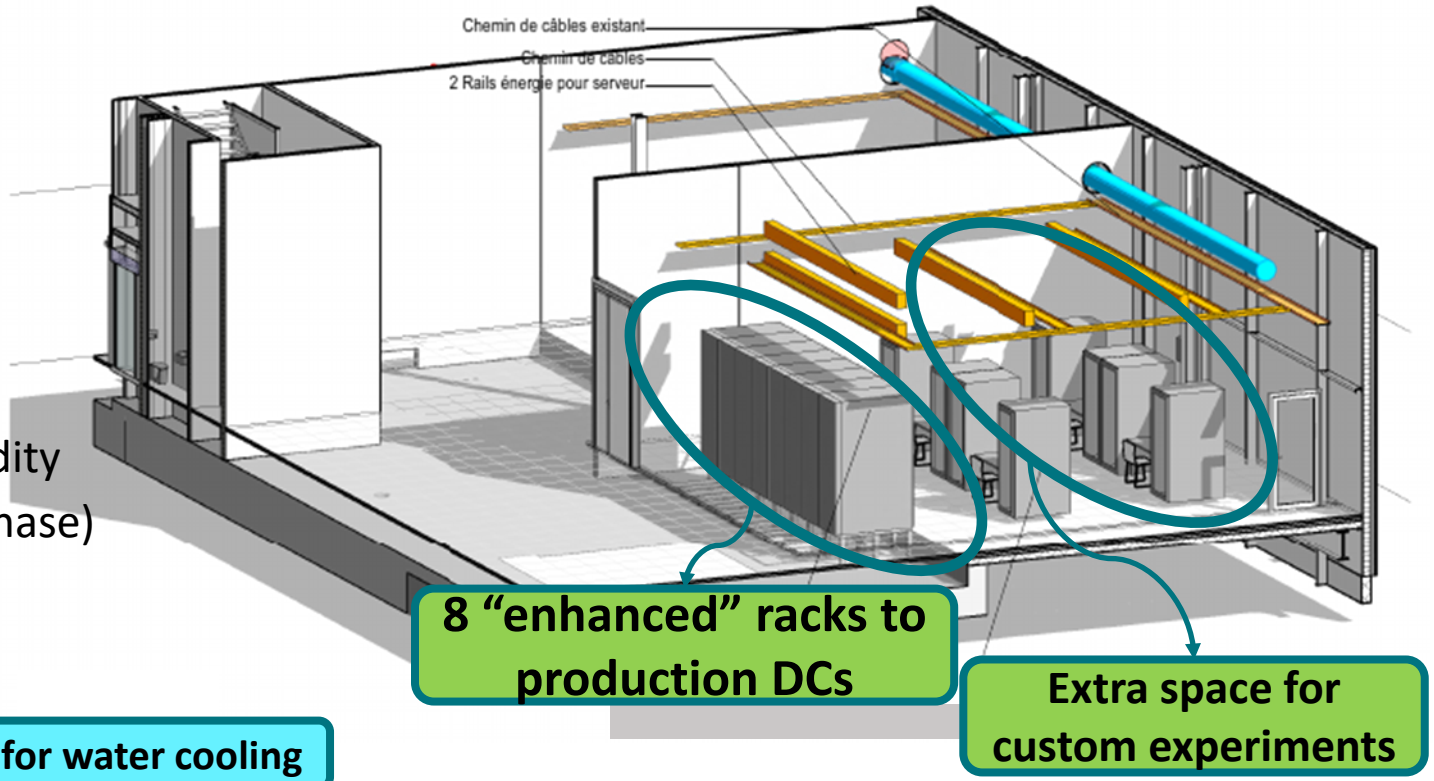
EcoCloud-EPFL Sustainability Experimental facility CCT

~150 m² of space for experiments on sustainable computing

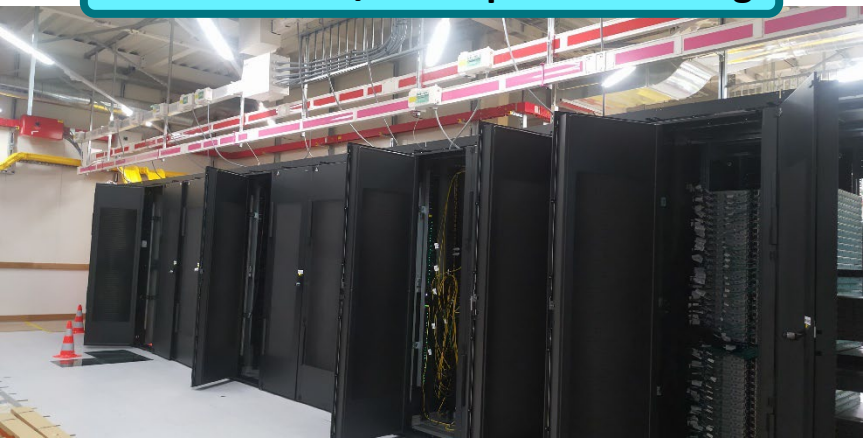
- Recycled racks/donations

Experimental support

- 50KW per rack/2.5m rack
- Monitoring: energy, temp., vibrat., humidity
- Cooling: air or water cooling (single-/2-phase)



Racks with air/water passive cooling

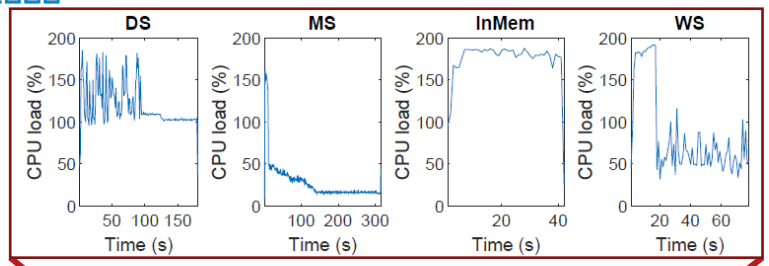


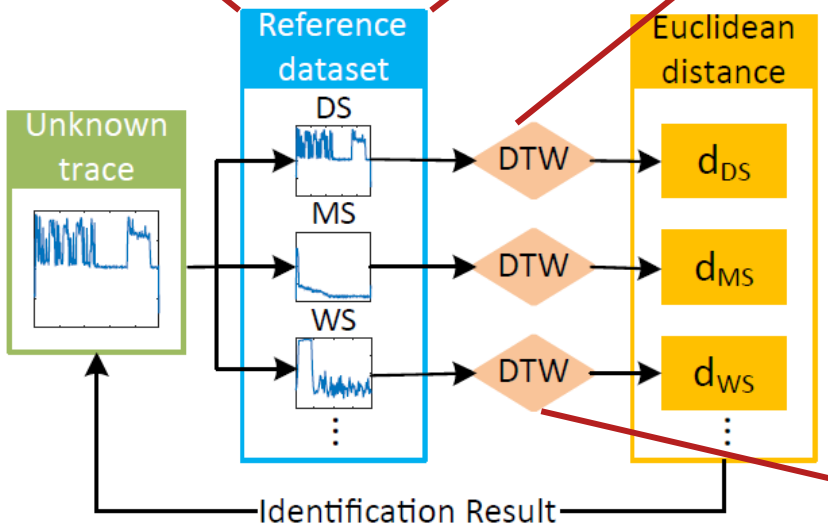
Pipes for water cooling

At EcoCloud-EPFL, we look forward to share this facility for new interesting projects!

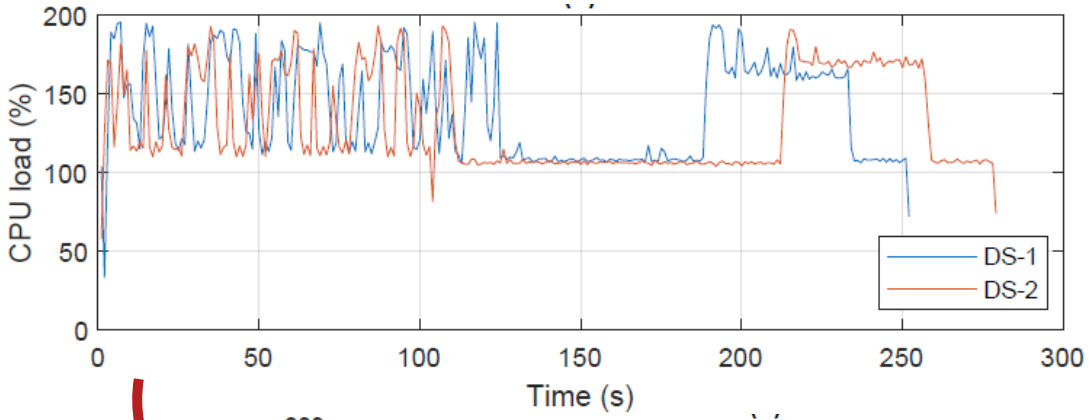
Application Identification

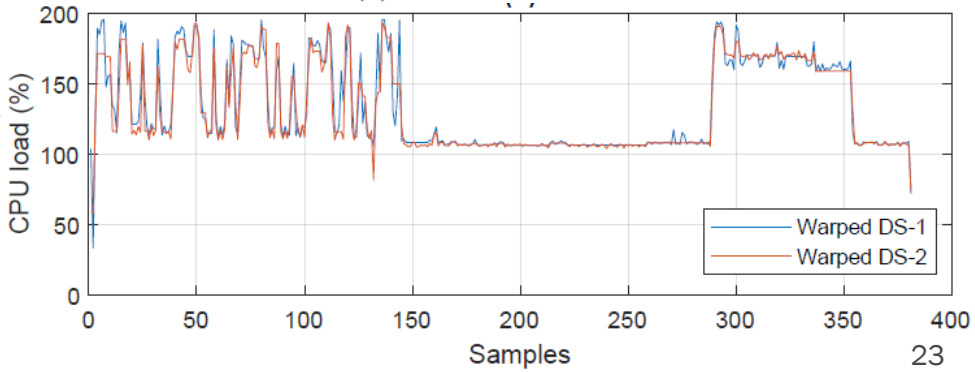
- Offline:
  - Create the reference dataset (Fingerprint), **no need for information from inside VM**
- Online:
  - Dyn. time warping (DTW) identification





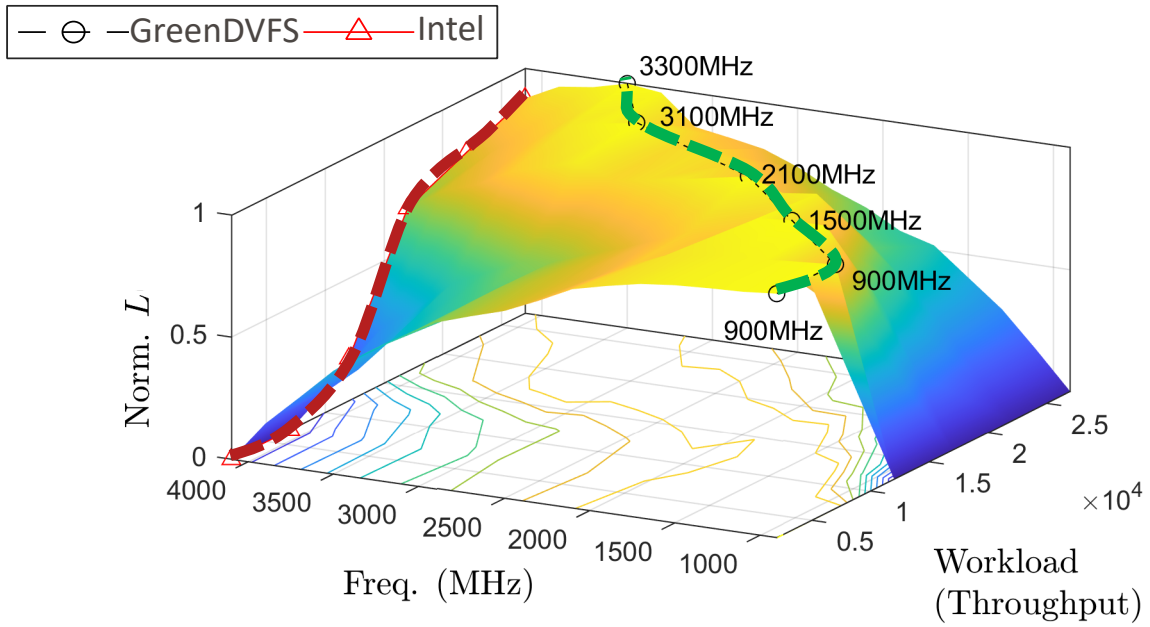
Dynamic time warping (DTW)



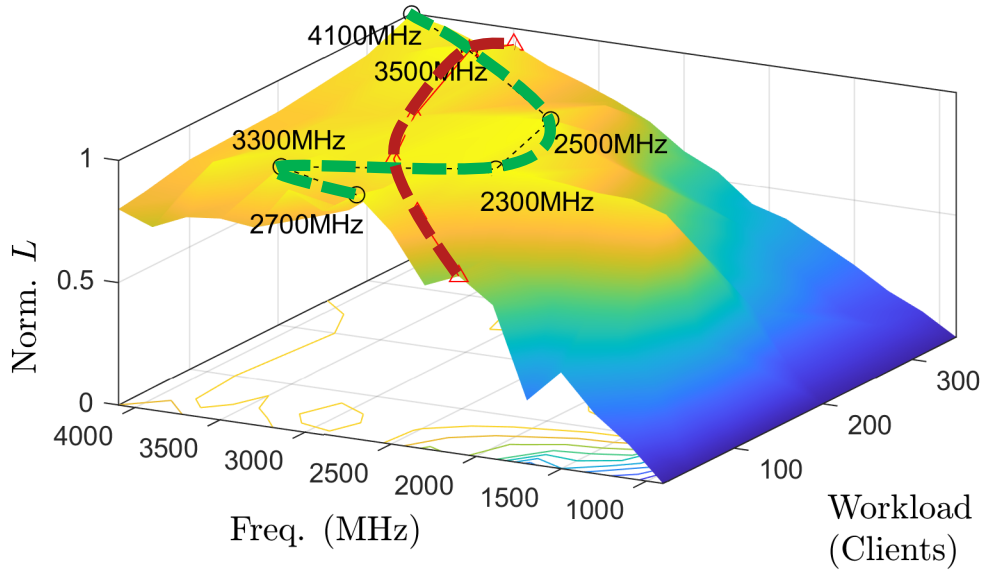


Key Outcome - Best Workload-Frequency Scaling Optimizer for Energy: Take it easy when going uphill!

- $L(f, wkl)$: Optimize performance, power, temperature
  - *intel-governor:* focus on CPU utilization



App: Data serving [3]



App: Web Search [3]

- GreenDVFS: $L(f, wkl) = \alpha \cdot Perf(f, wkl) - \beta \cdot Power(f, wkl) - \gamma \cdot Temp(f, wkl)$

$\alpha = 0.5, \beta = 0.3, \gamma = 0.2$

But Linux/Proprietary **Scaling Governors Are Not Optimal**

- *powersave:* lowest frequency

- *performance:* highest frequency

- (default) *schedutil* or *intel-governor*:
  - Frequency scaling with CPU utilization

- Take home messages:

powersave is the most energy-intensive!

Linux/propr. scaling governors are sub-optimal!