

Sustainable cloud and Edge for Emerging Workloads



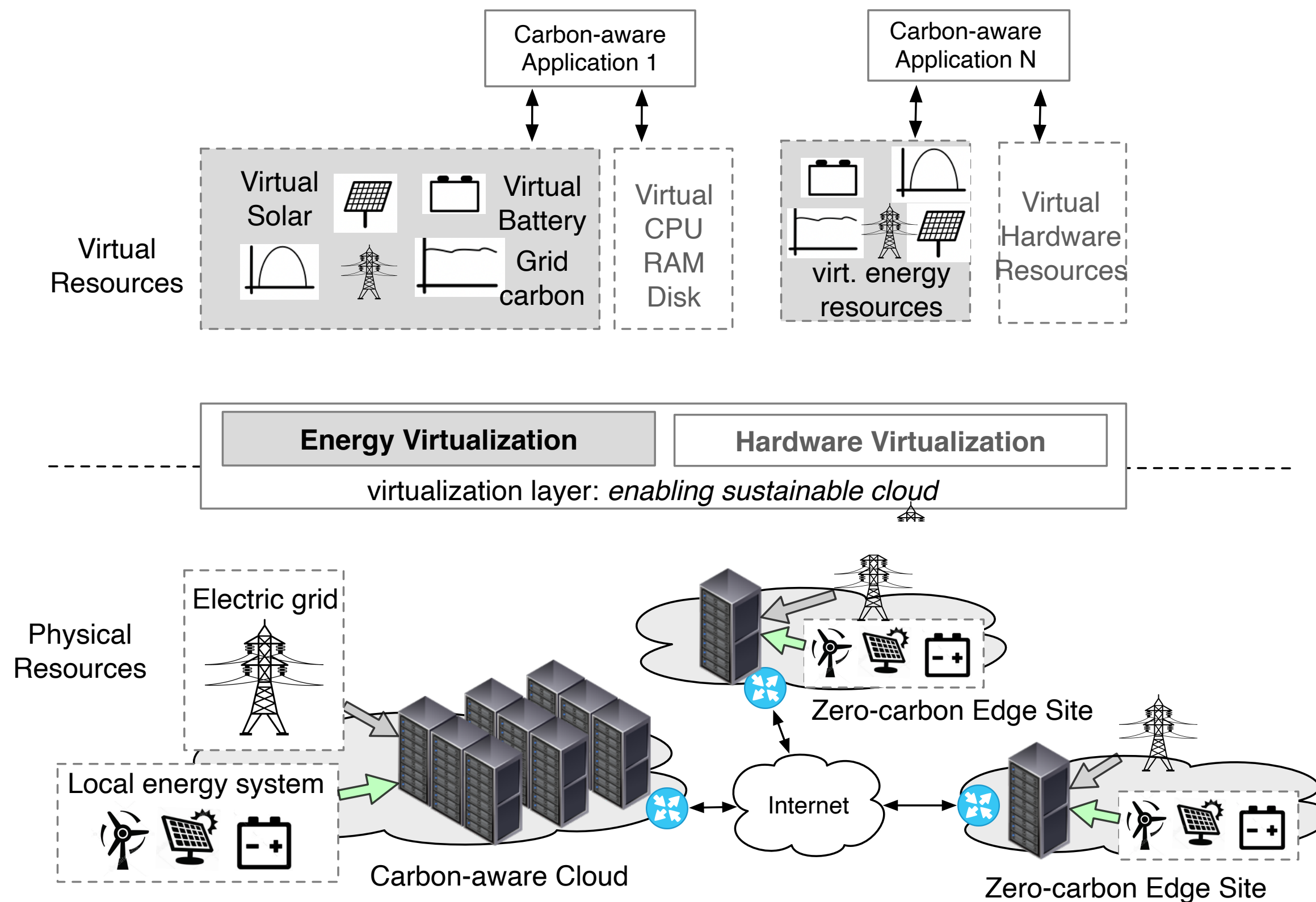
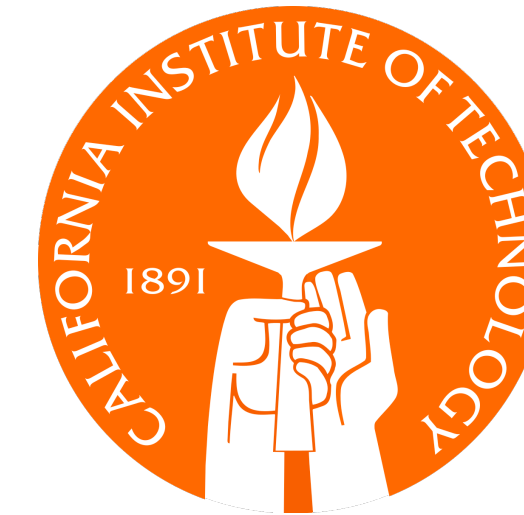
Tian Guo

Associate Professor

tian@wpi.edu



Sustainable Cloud-Edge Infrastructure Vision



End-to-end approach that delegates **resource management** to the **applications**

Energy-Efficiency **Cannot** Mitigate Energy Growth

> Few opportunities left to optimize energy-efficiency

- Many cloud data centers operate near peak efficiency

Moving forward:

Exponential cloud growth ->
Exponential energy demand

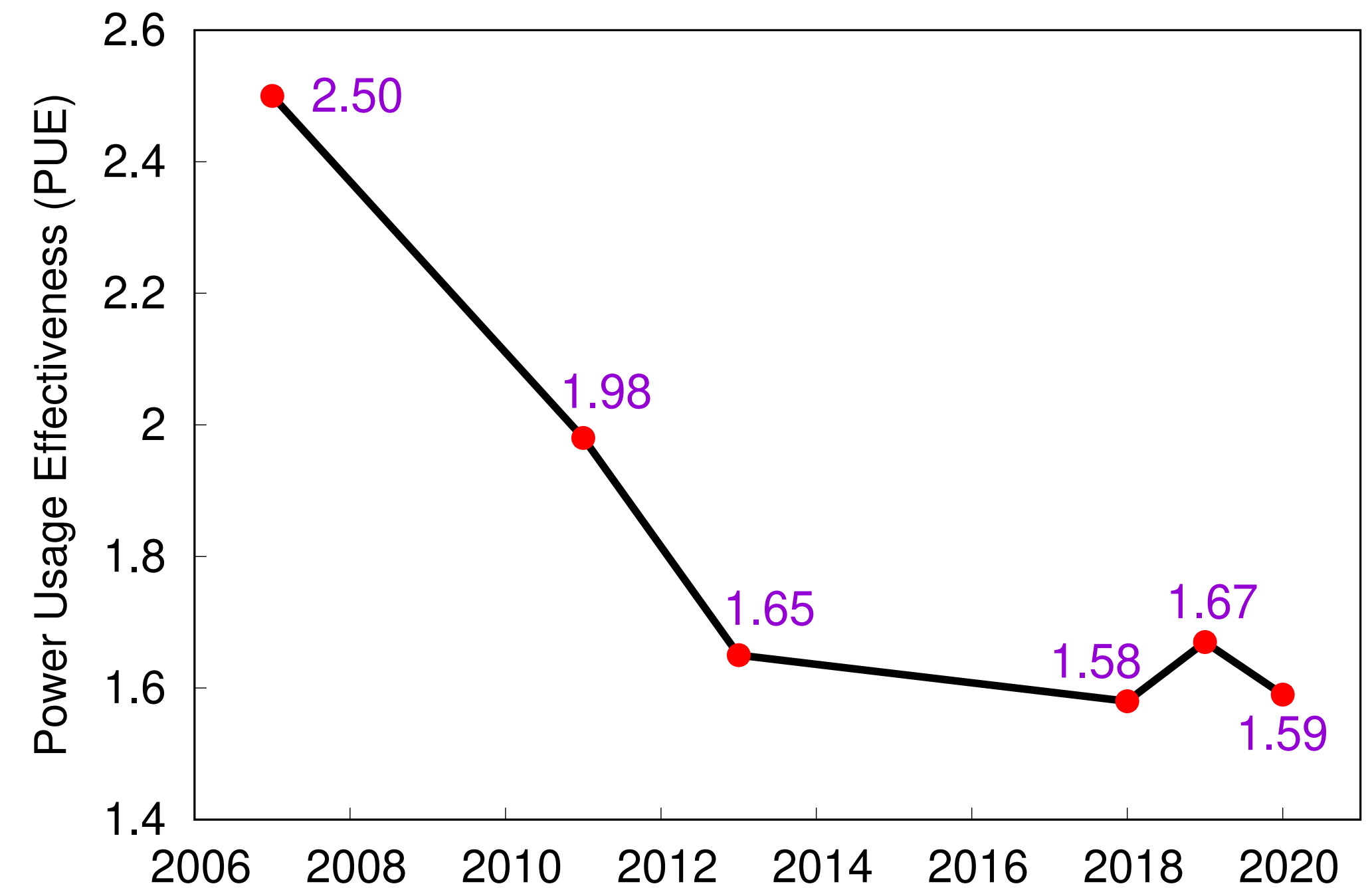
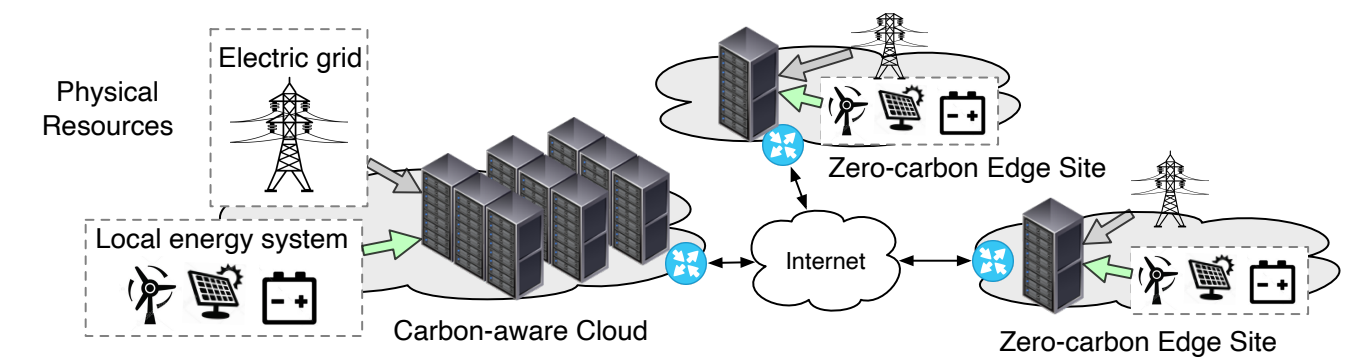
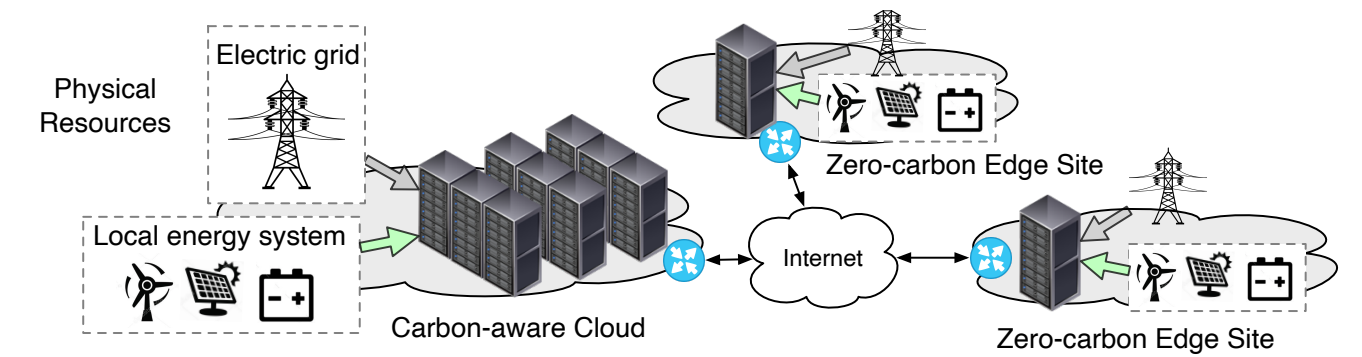


Figure source: Data center PUEs have been flat since 2013, Uptime Institute

Cloud's Energy Usage is Not the Problem

- > Optimizing carbon differs from optimizing energy
- **Energy-efficiency** – computations per unit energy consumed
- **Carbon-efficiency** – computations per unit carbon emitted



Energy-**inefficient**
Carbon-efficient



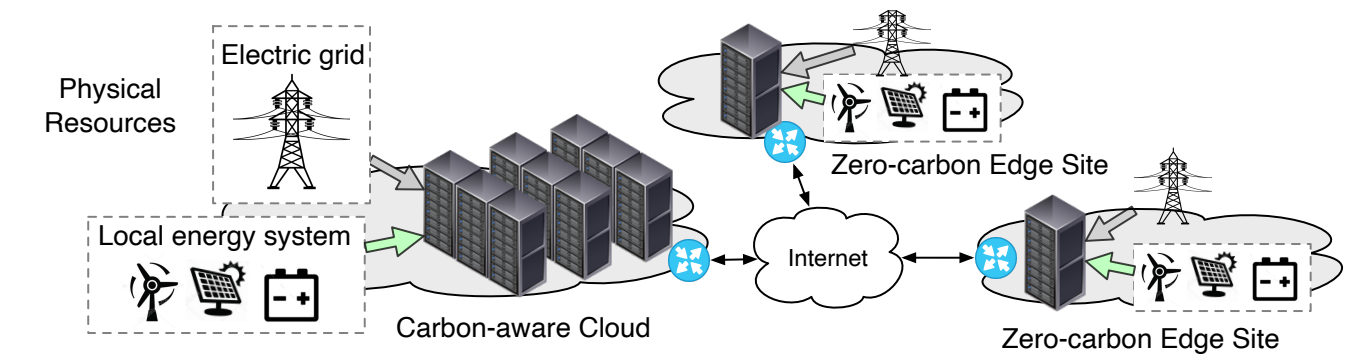
Energy-efficient
Carbon-**inefficient**

Prior work:

Decades of work on **energy**-efficiency
Very little work on **carbon**-efficiency

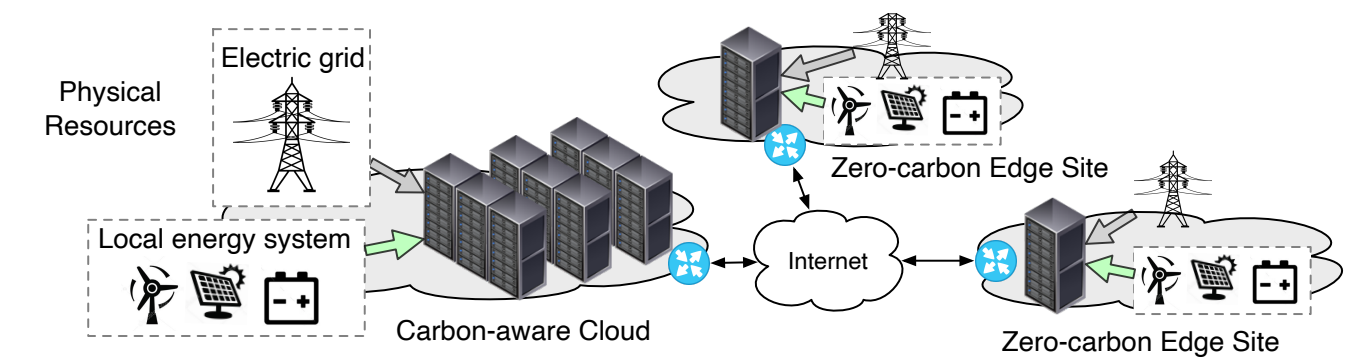
A Carbon-First Approach to Designing Clouds

- > Elevate carbon to be a **first-class** metric
 - Optimize **carbon-efficiency** instead of energy-efficiency
 - Develop approaches that directly reduce carbon

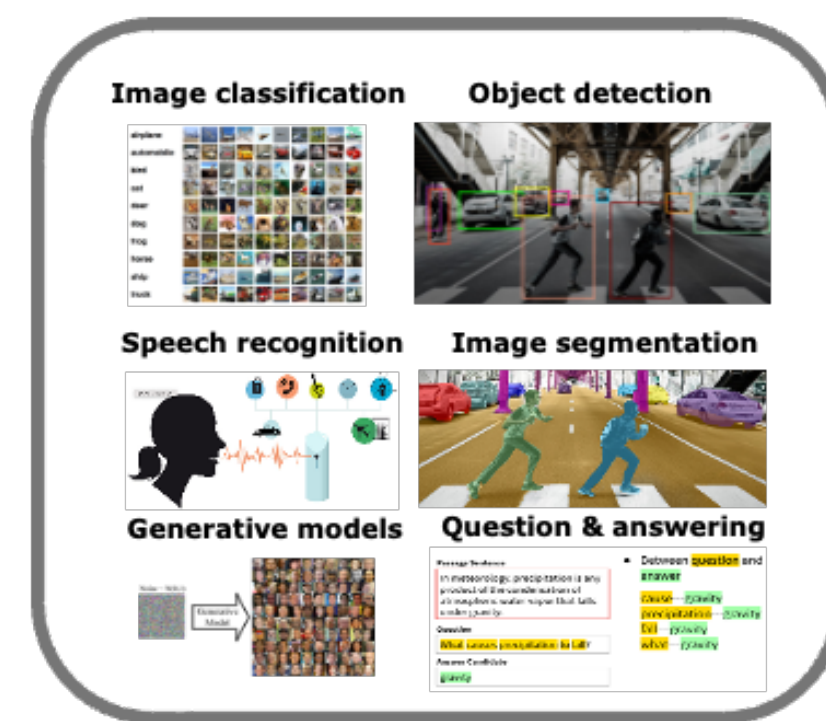


A Carbon-First Approach to Designing Clouds

- > Elevate carbon to be a **first-class** metric
 - Optimize **carbon-efficiency** instead of energy-efficiency
 - Develop approaches that directly reduce carbon



Challenges



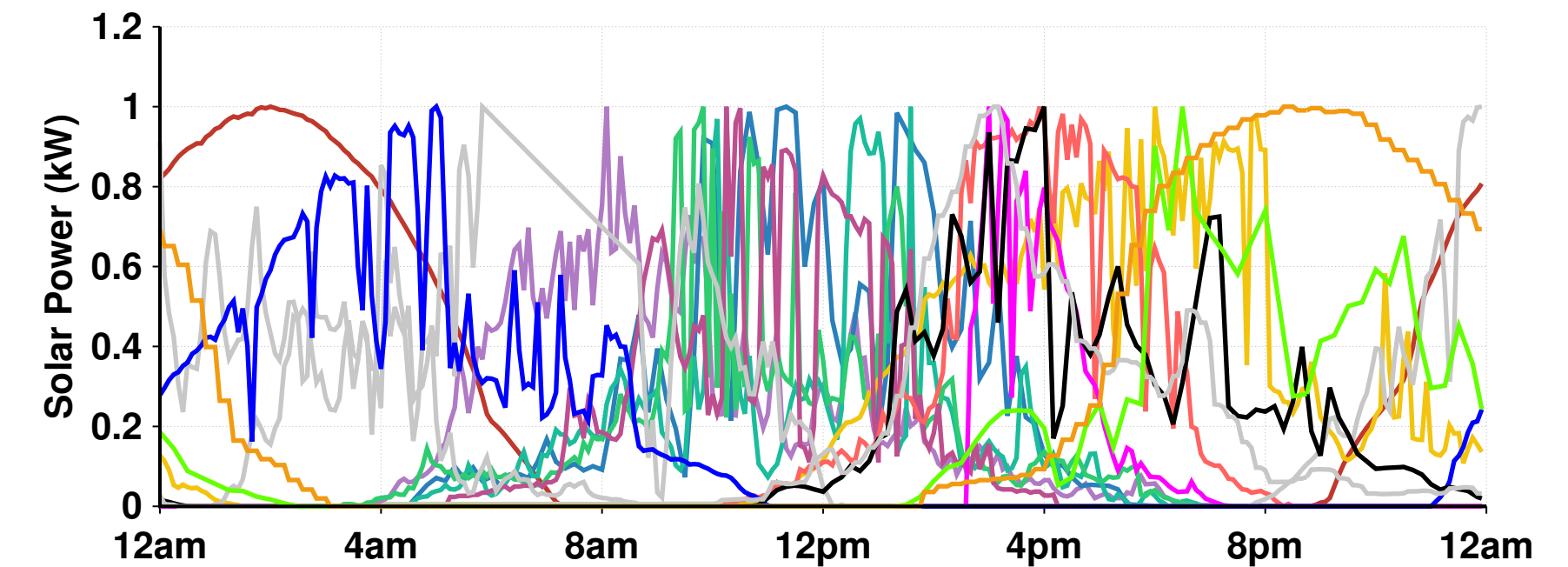
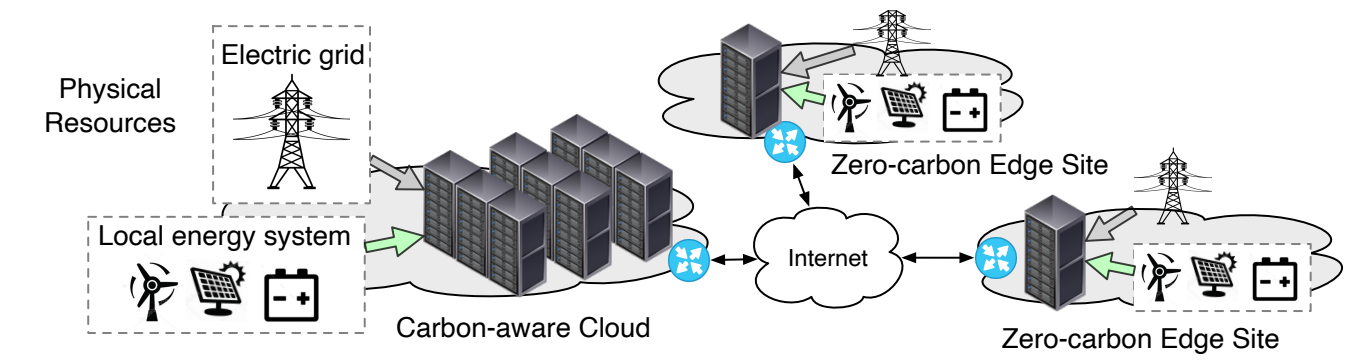
Diverse applications

Energy

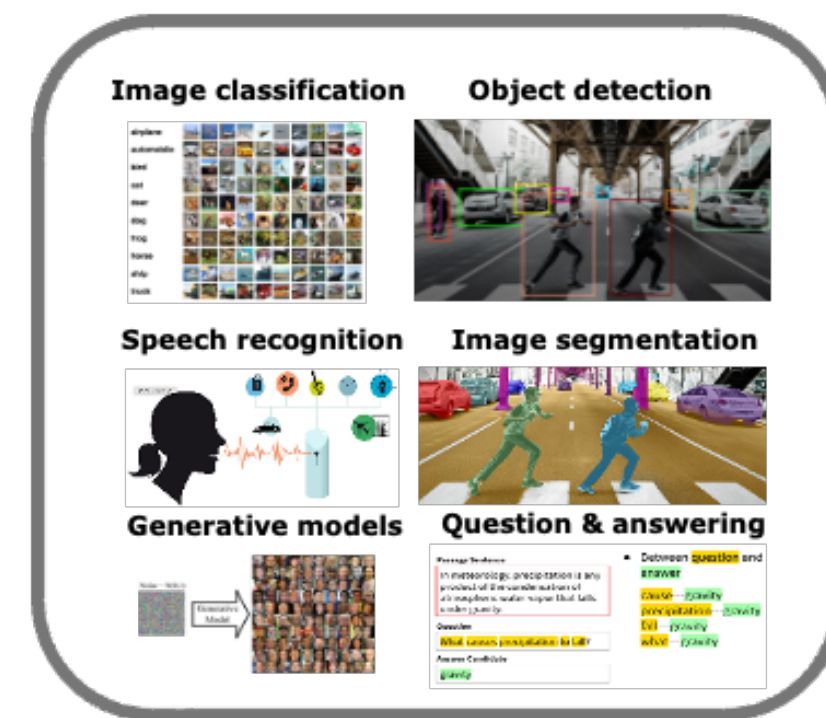
Continuous and reliable power supply

A Carbon-First Approach to Designing Clouds

- > Elevate carbon to be a **first-class** metric
 - Optimize **carbon-efficiency** instead of energy-efficiency
 - Develop approaches that directly reduce carbon



Challenges



Energy

Continuous and reliable power supply

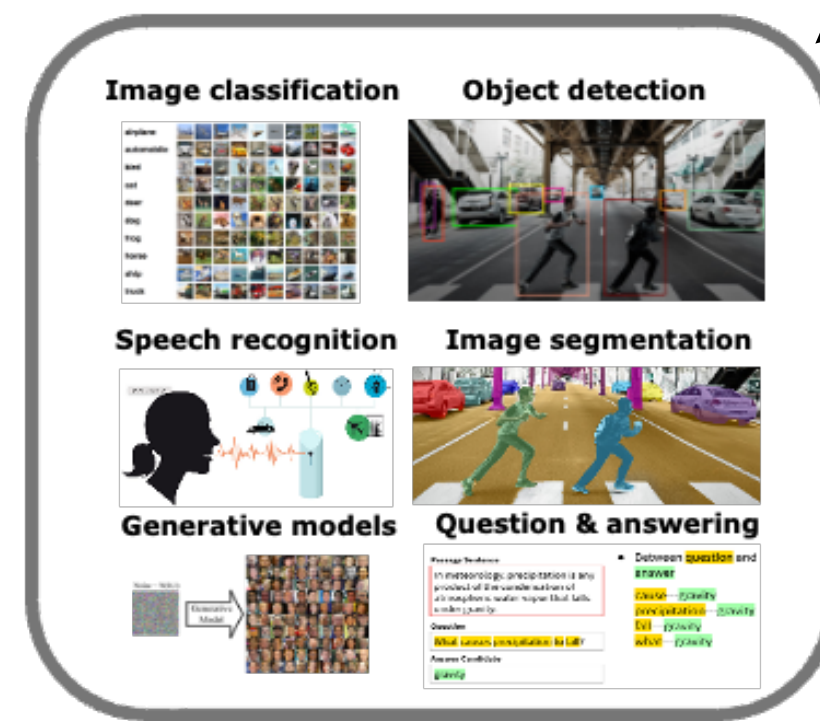
Diverse applications

A Carbon-First Approach to Designing Clouds

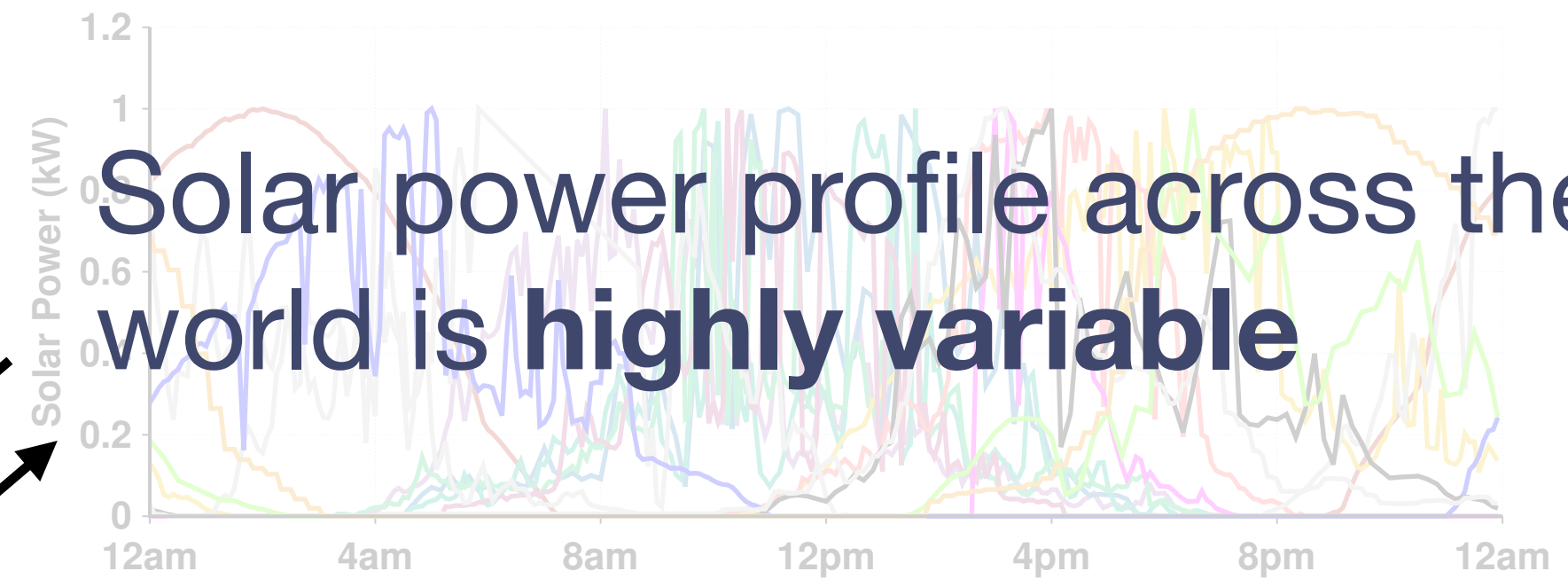
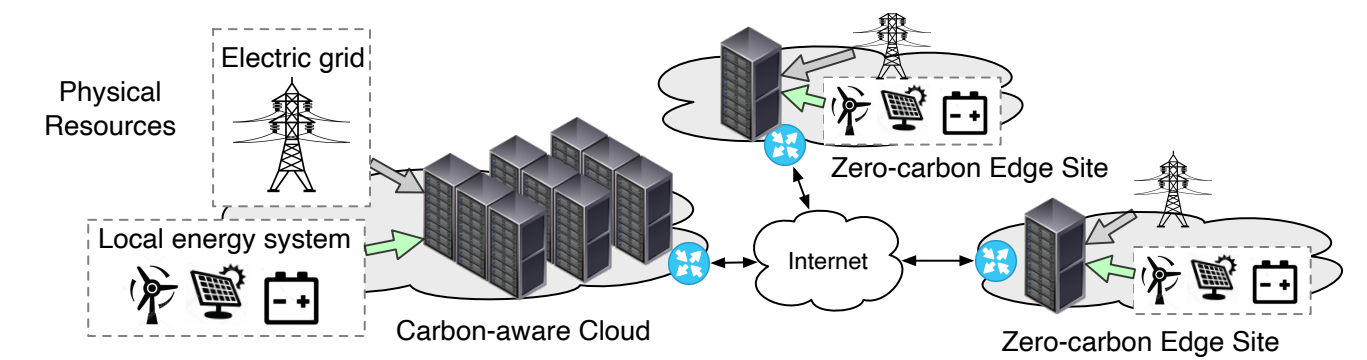
> Elevate carbon to be a **first-class** metric

- Optimize **carbon-efficiency** instead of energy-efficiency
- Develop approaches that directly reduce carbon

Challenges



Diverse applications



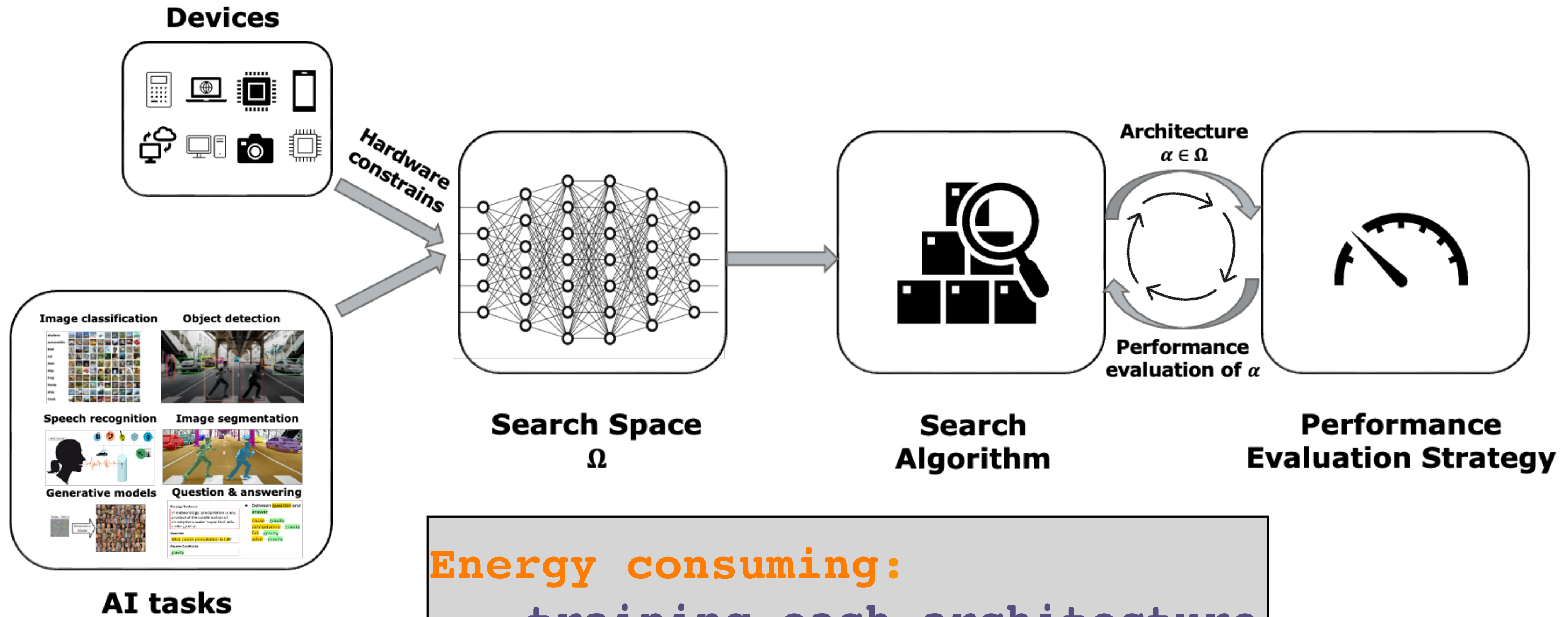
Solar power profile across the world is **highly variable**

Visibility
Control

Energy



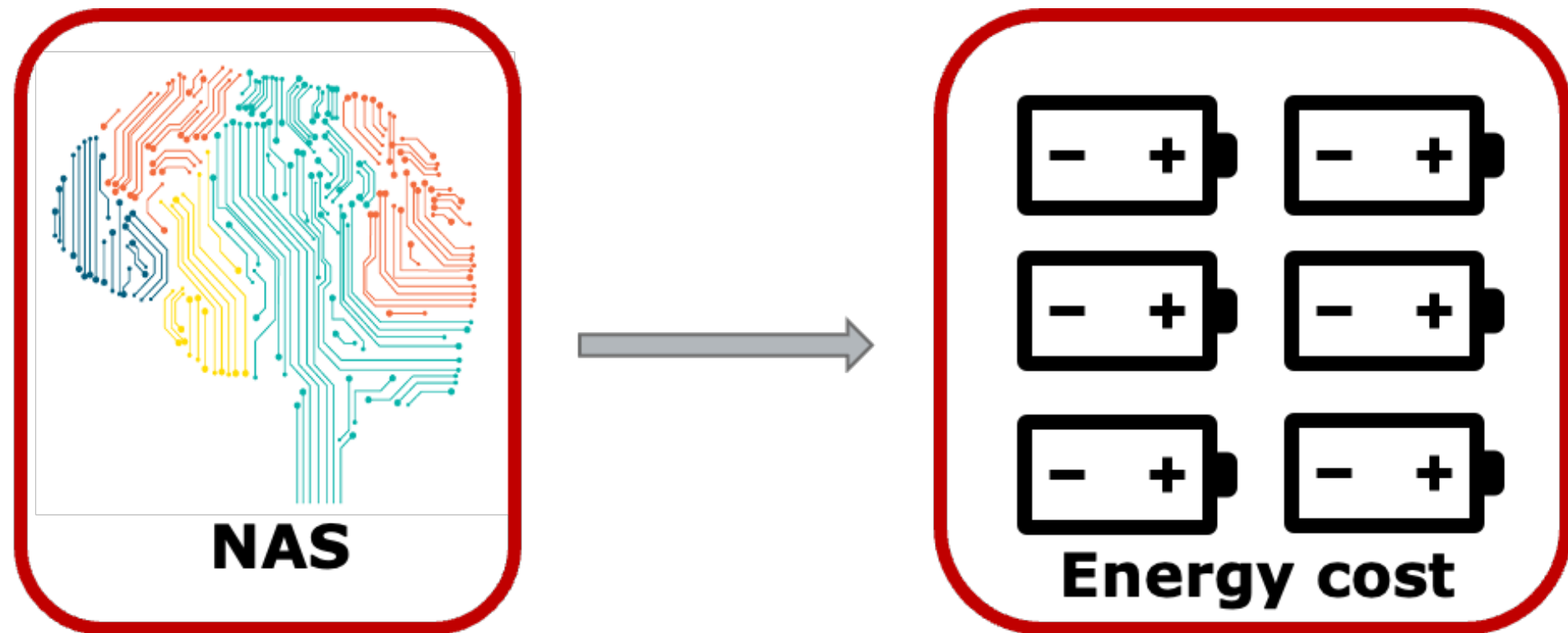
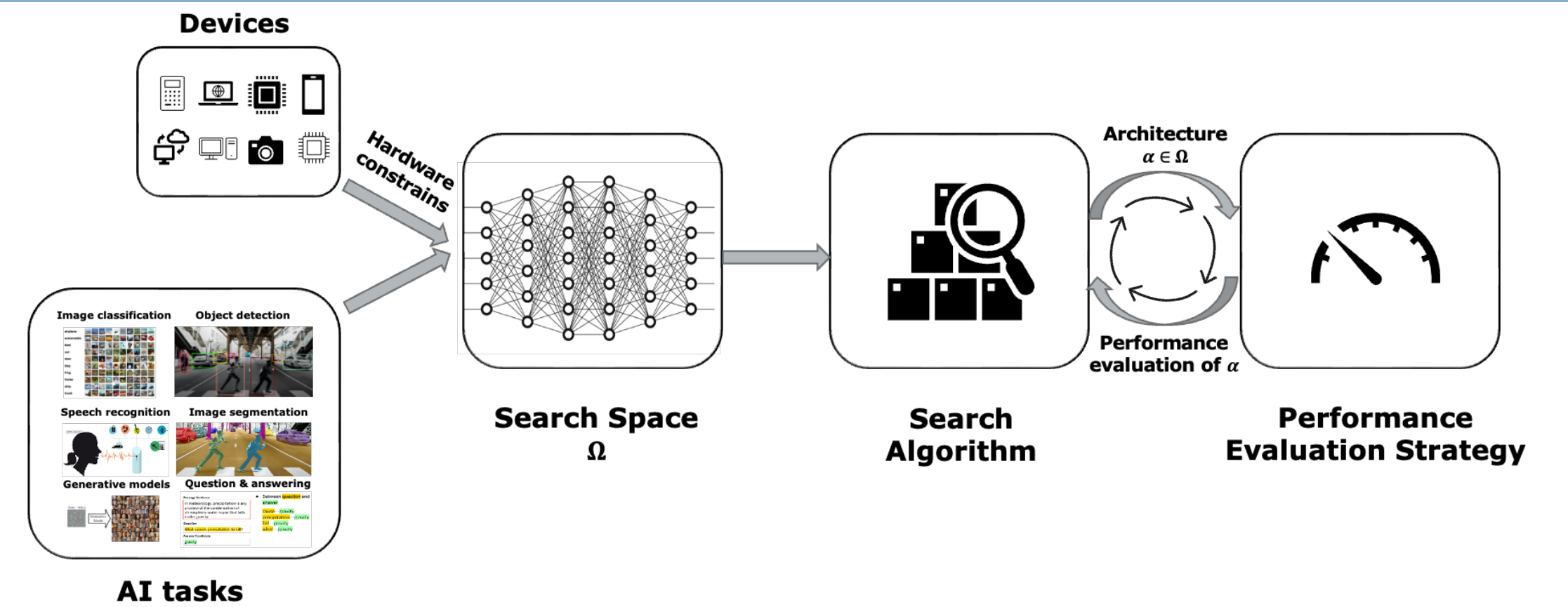
Neural Architecture Search (NAS)



Energy consuming:

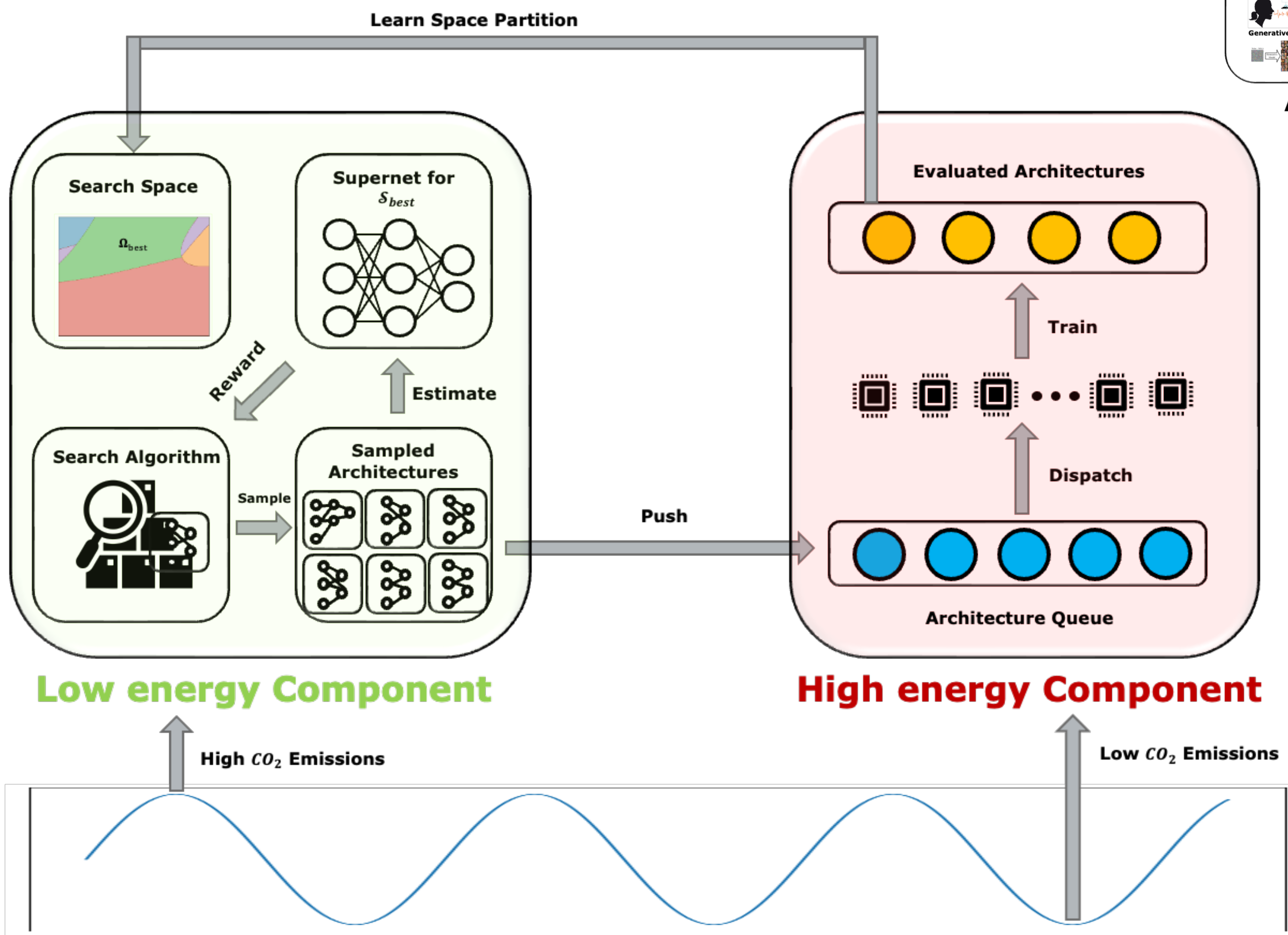
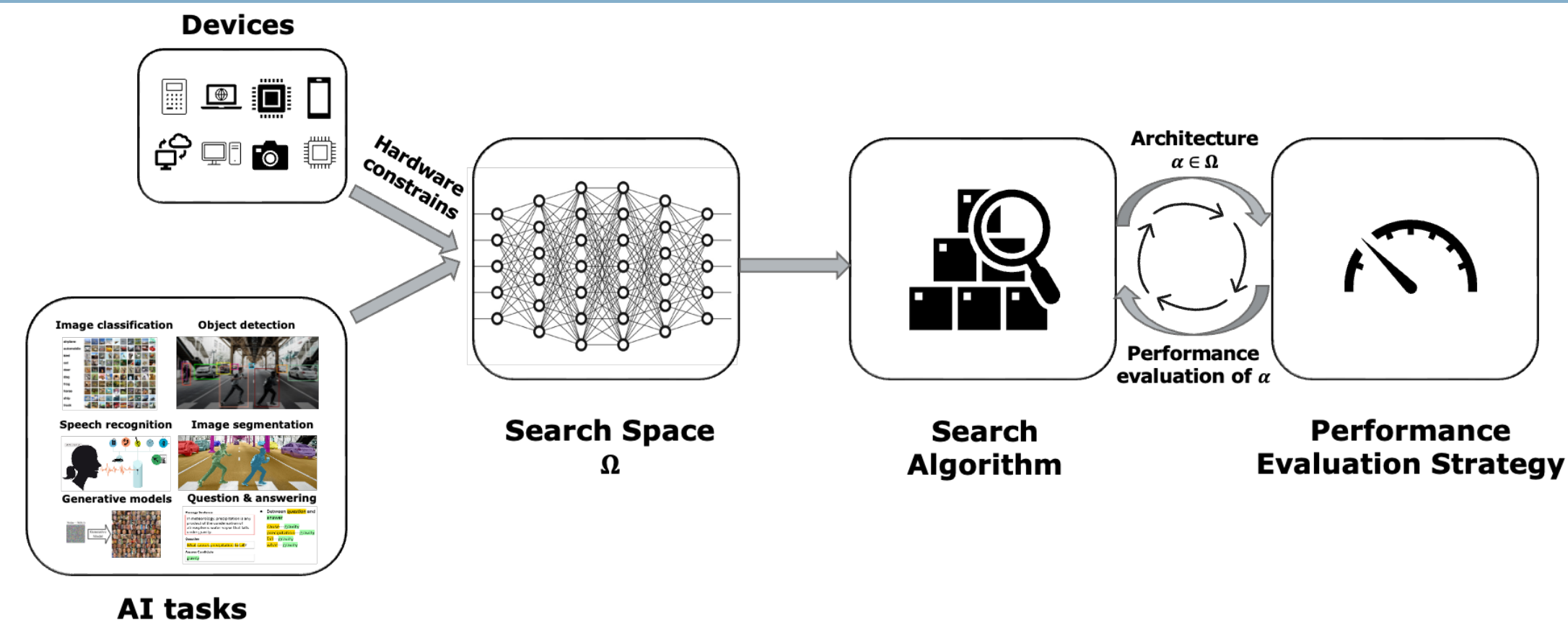
- training each architecture can take hours
- might need to train hundreds to find good ones

Opportunities



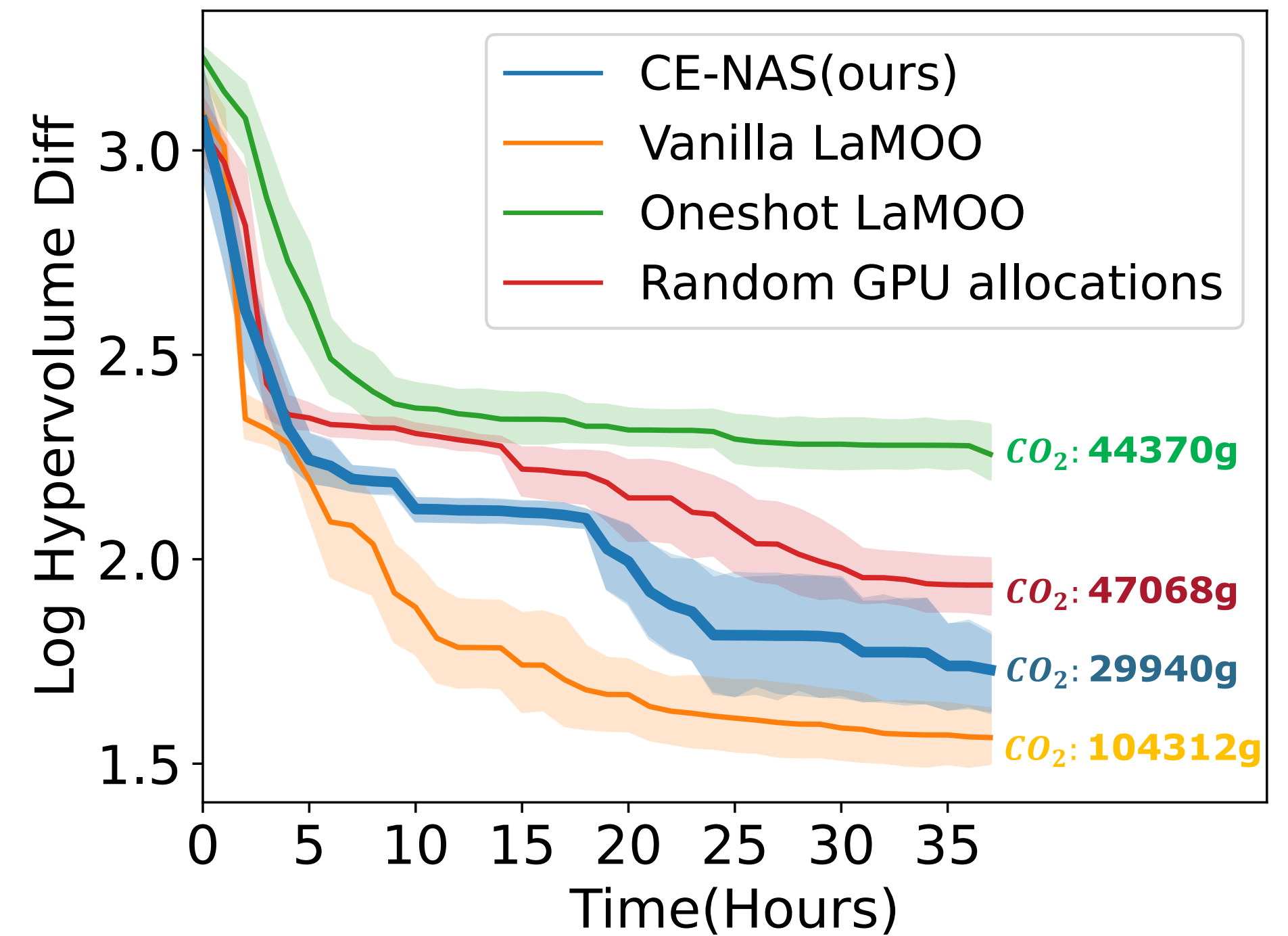
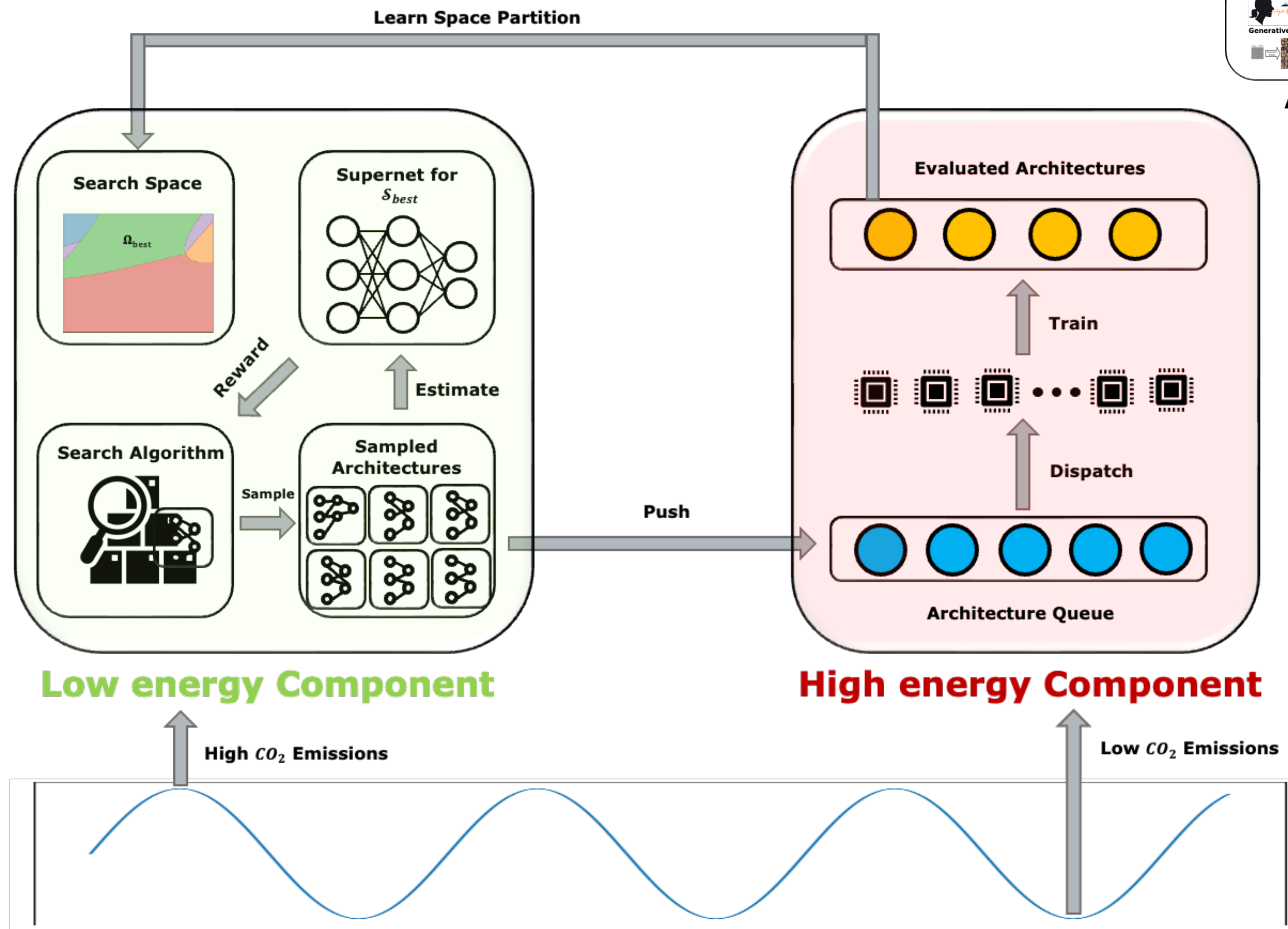
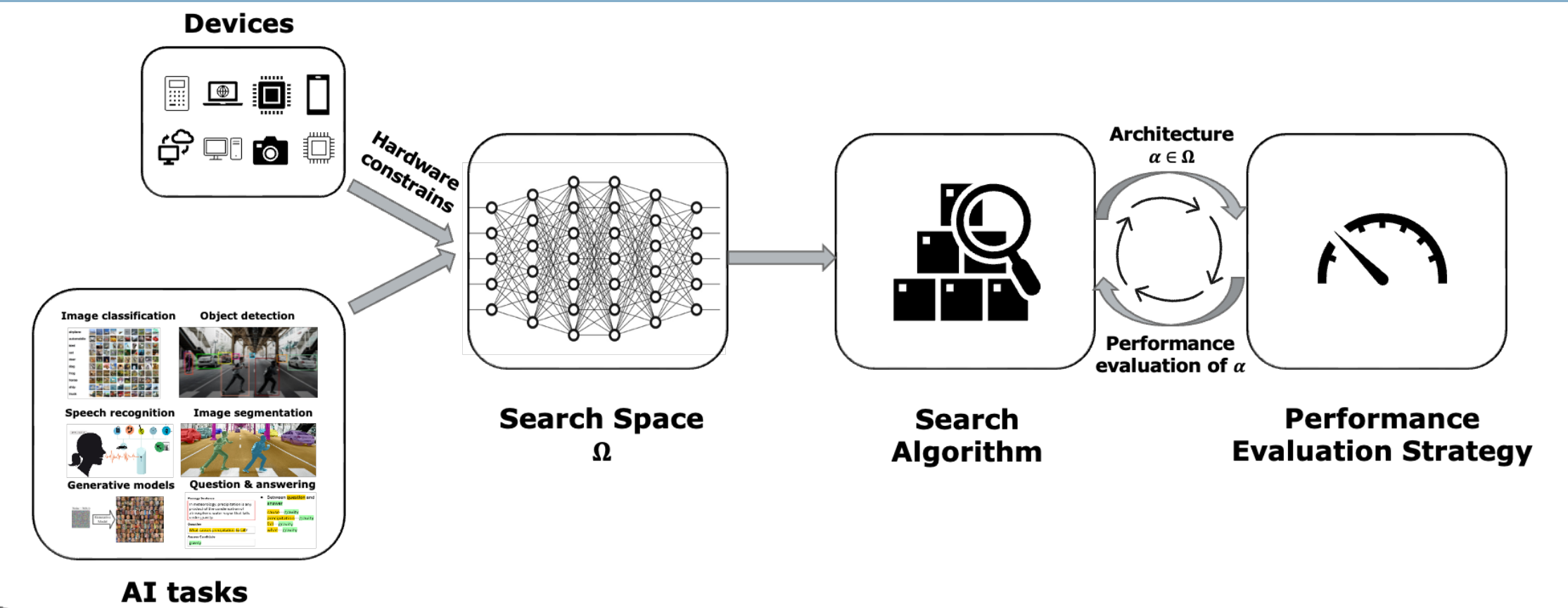
- > Not all **NAS steps** are equally energy consuming
- > Different **NAS strategies** also have different energy requirements

CE-NAS Key Idea



- Goals**
 - Less carbon emissions
 - High search efficiency
 - SoTA architectures
- Baselines**
 - Few-shot NAS
 - LaMOO NAS

Preliminary Result



Our CE-NAS has the lowest relative carbon emission while achieving the second best HV_{\log_diff}