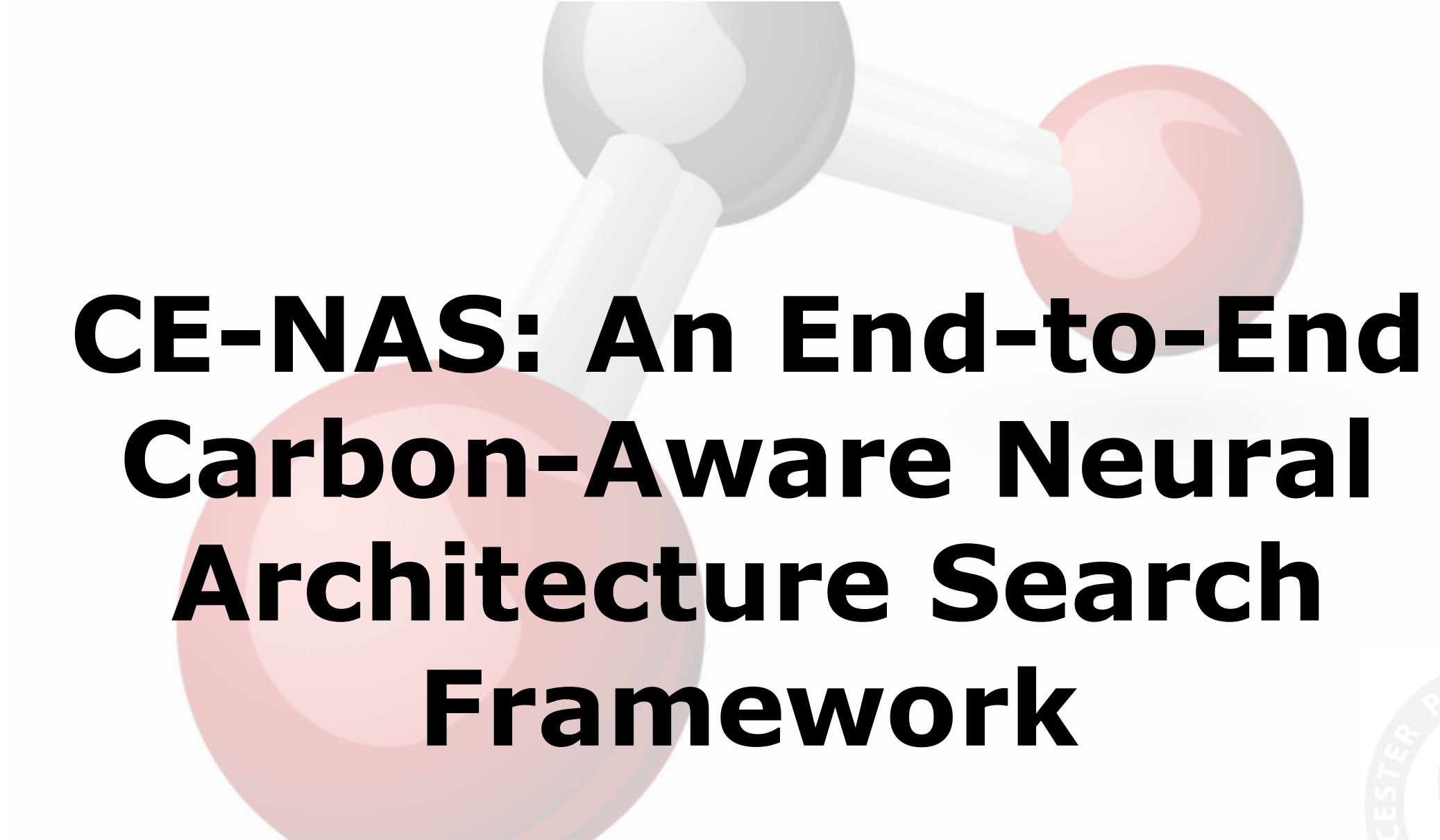
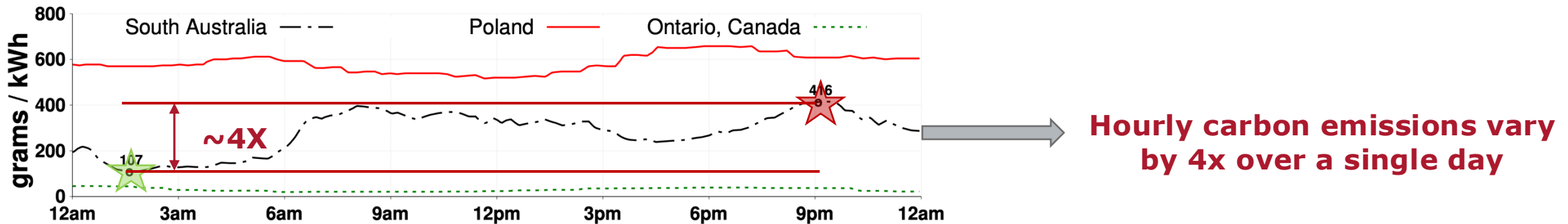
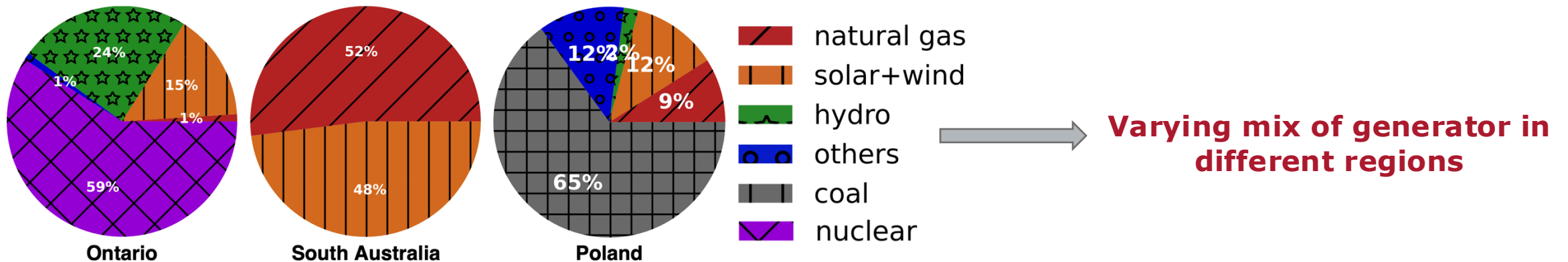




CE-NAS: An End-to-End Carbon-Aware Neural Architecture Search Framework

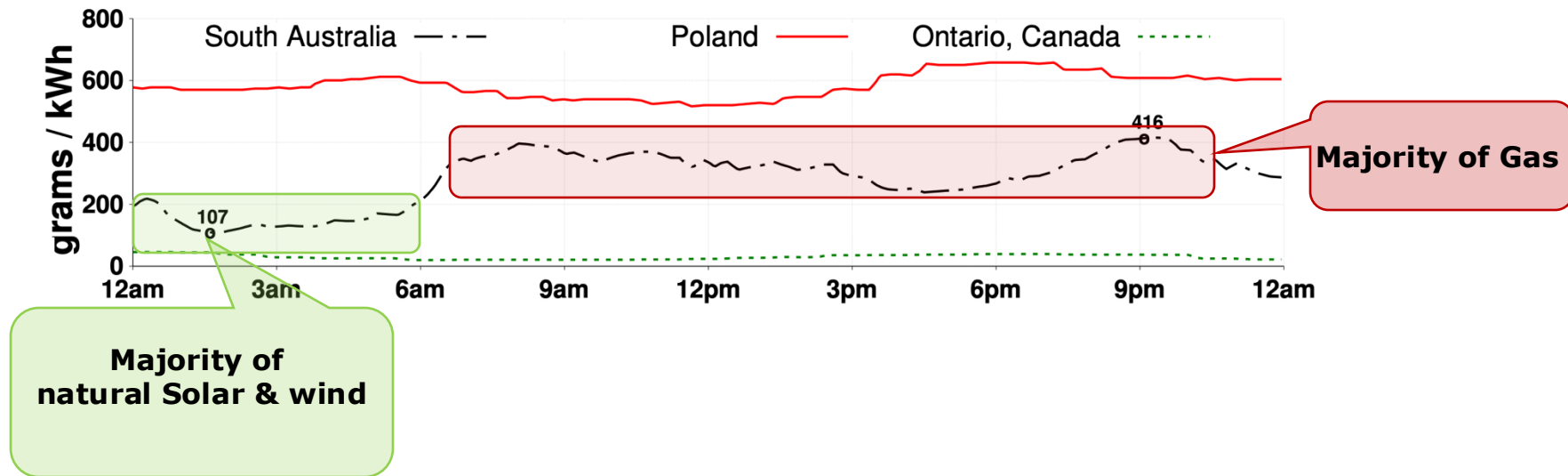


Statistics of Carbon Emissions¹

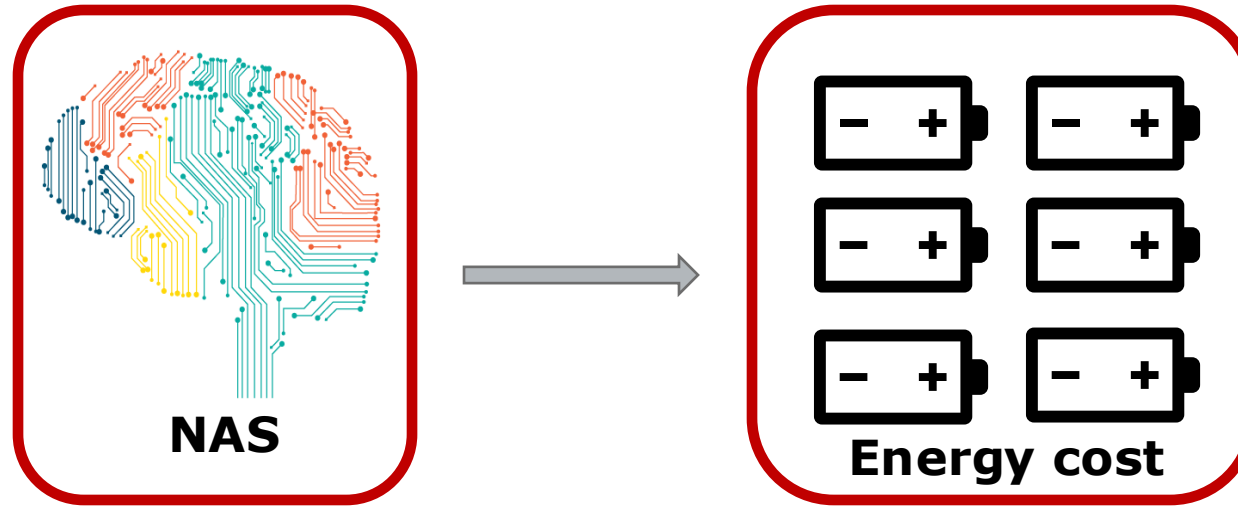


1. Enabling Sustainable Clouds: The Case for Virtualizing the Energy System. Noman Bashir, Tian Guo, Mohammad Hajiesmaili, David Irwin, Prashant Shenoy, Ramesh Sitaraman, Abel Souza, Adam Wieman. ACM Symposium on Cloud Computing 2021.

Statistics of Carbon Emissions¹



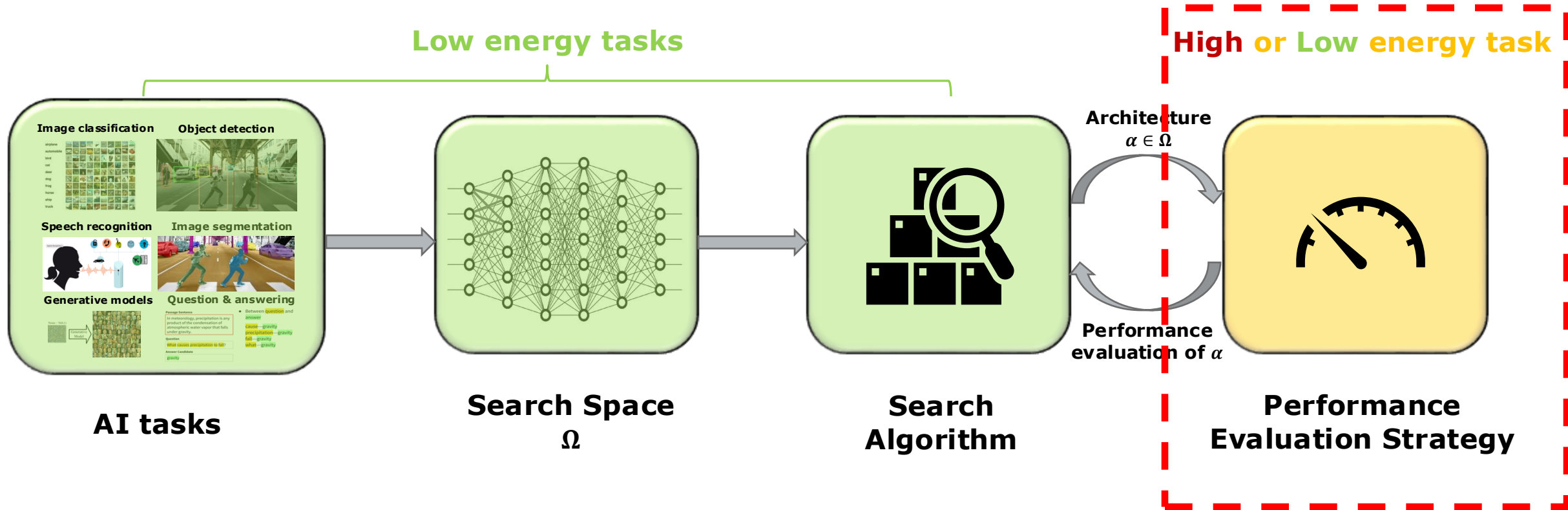
Problem Statement



NAS can be energy consuming

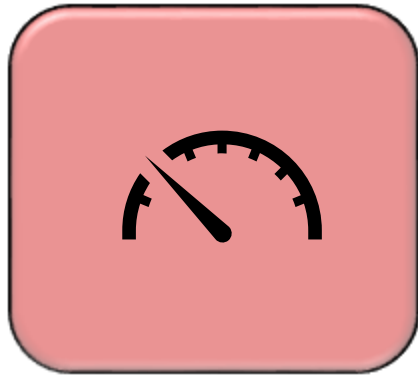
But not all NAS steps are equally energy consuming, and different NAS strategies will also have different energy requirements

Task Energy Requirement Disparity

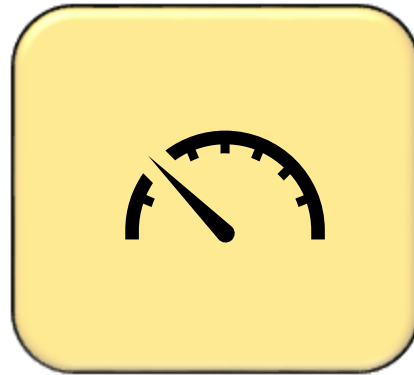


Evaluation Energy Requirement Disparity

High energy task



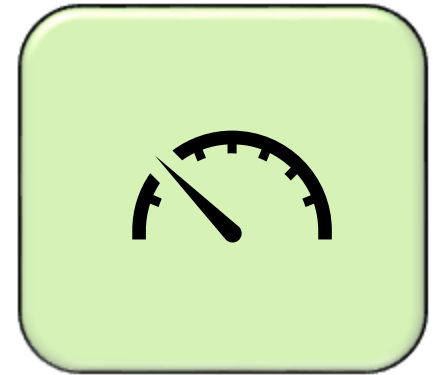
**Vanilla NAS
Evaluation**



**Performance
Evaluation Strategy**

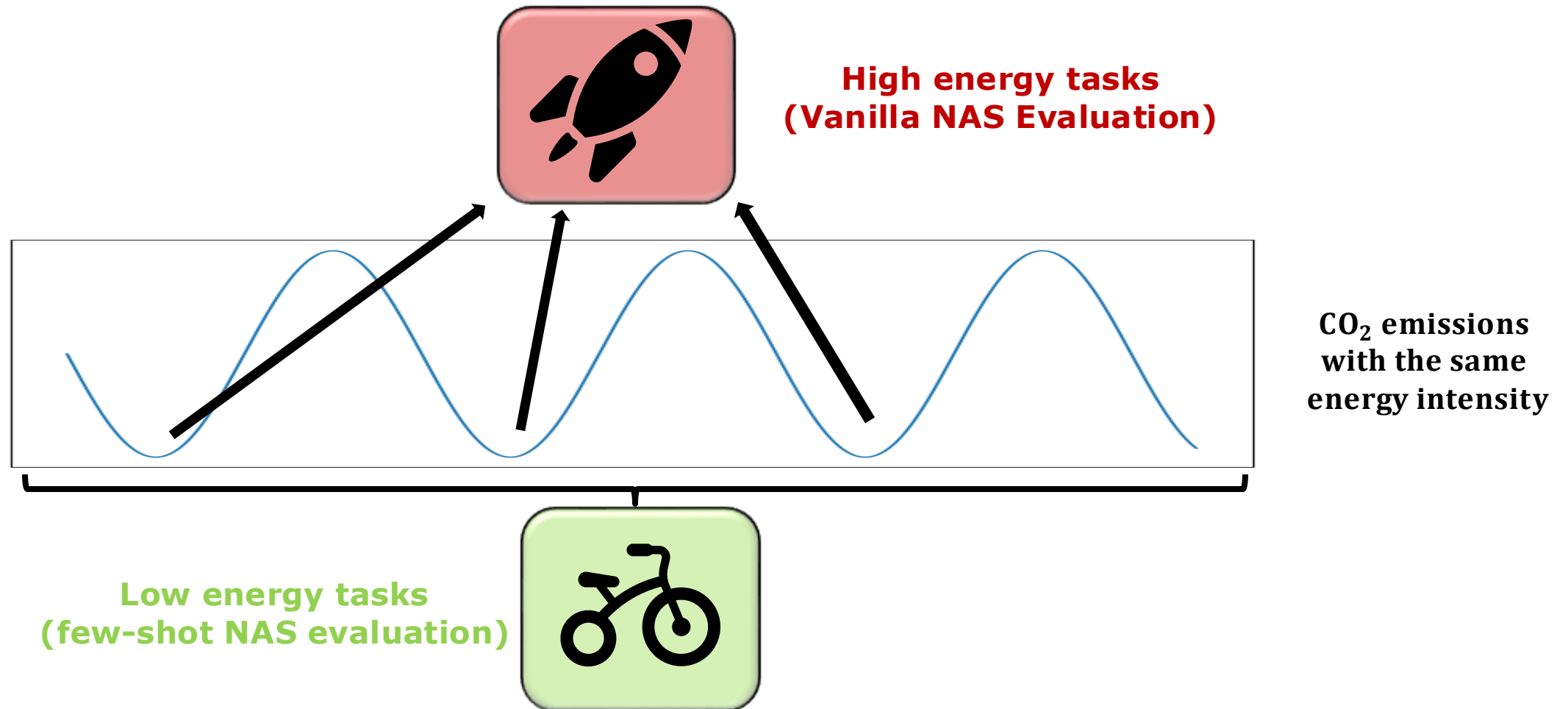


Low energy task

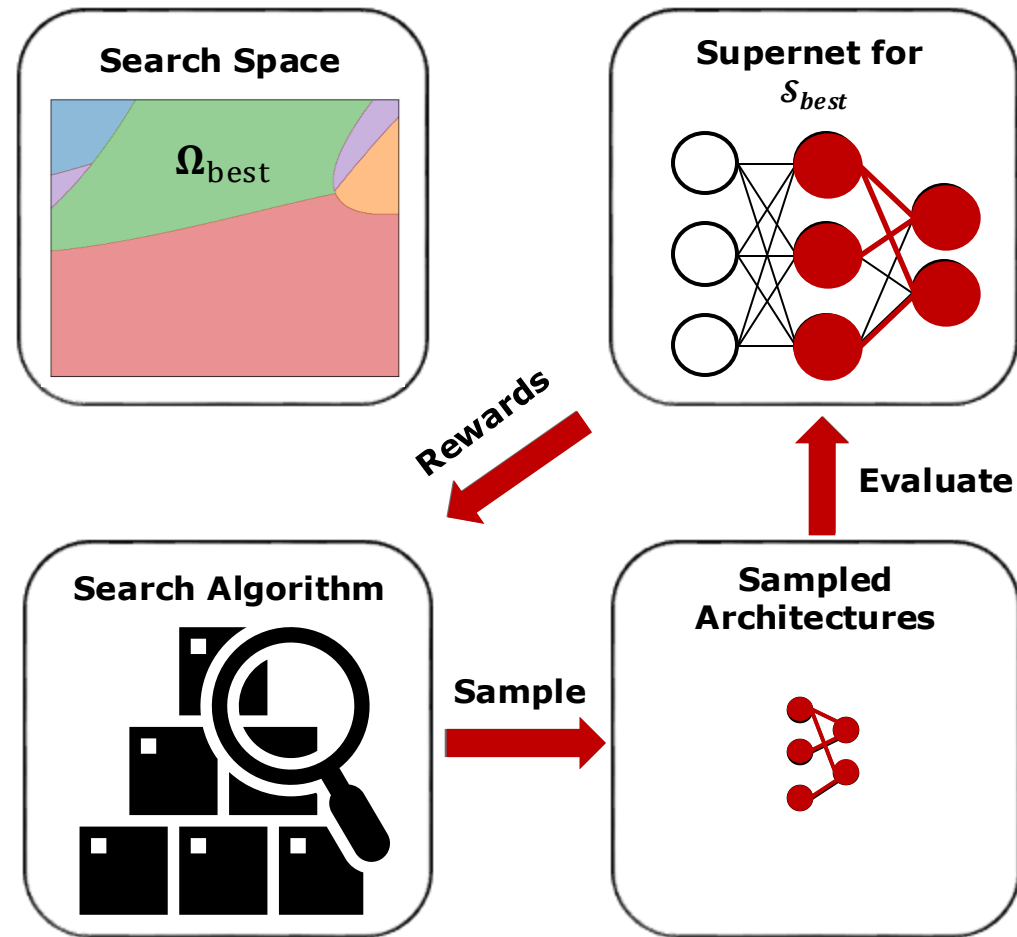


**One/few-shot NAS
Evaluation**

Our Key Idea



Low Energy Component

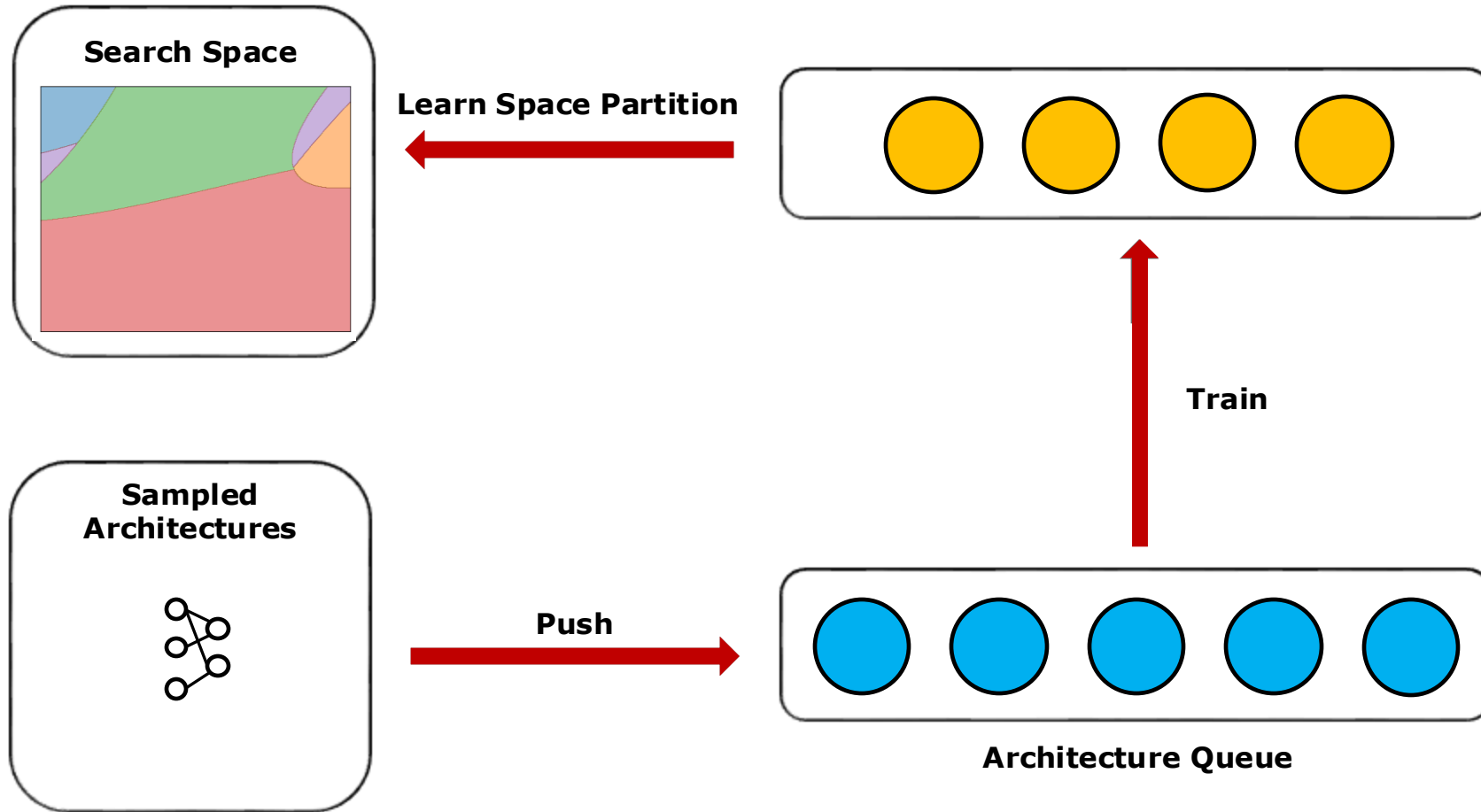


Low energy component in our NAS framework, will run during the high carbon periods

Low energy Component

Worcester Polytechnic Institute

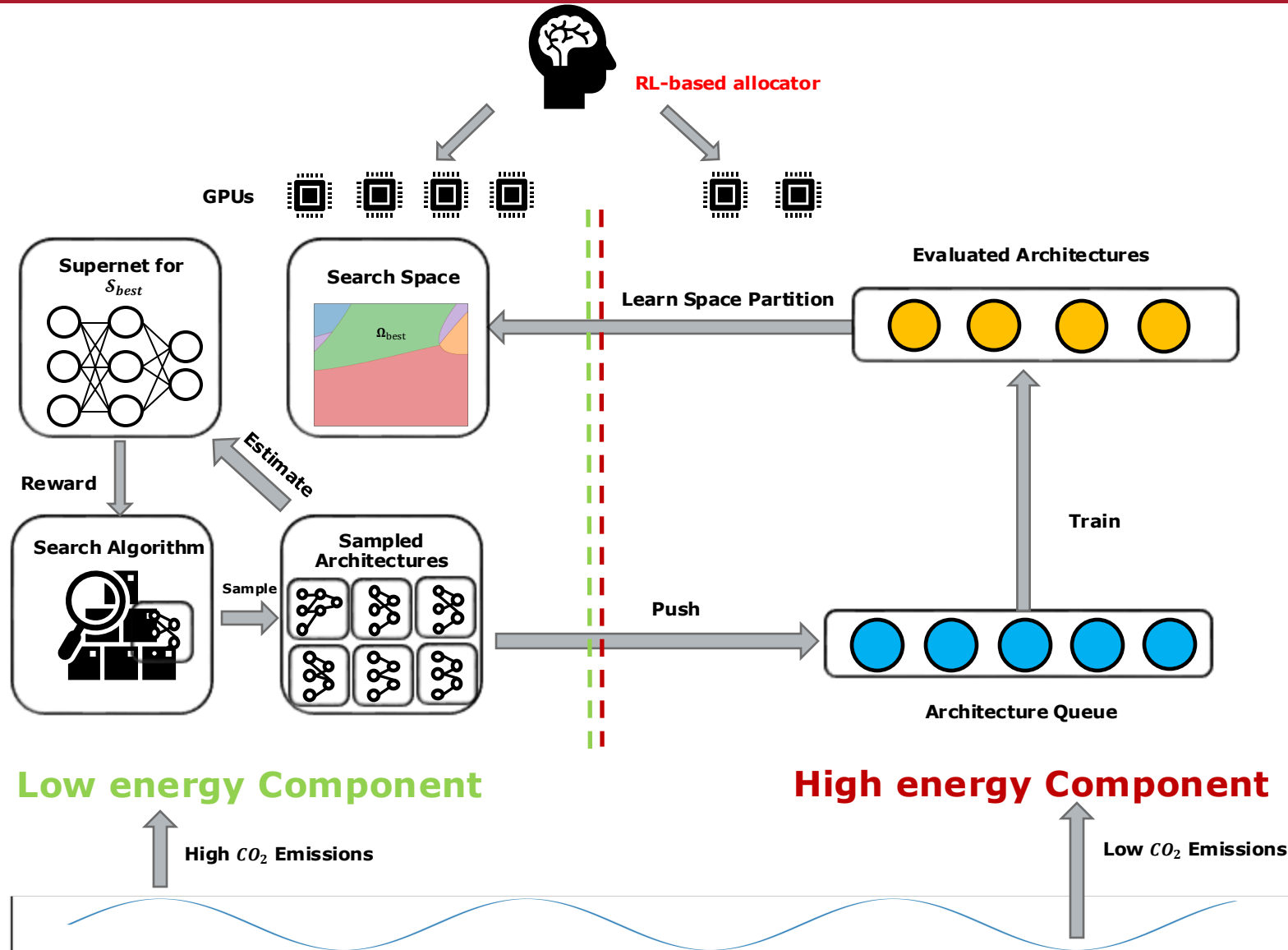
High Energy Component



Multi-objective
Optimization by
Learning Space
Partitions

**High energy component
in our NAS framework,
will run during the low
carbon periods**

How to allocate GPU resource



Reinforcement learning(RL)-based allocation method

Reinforcement Learning -- State



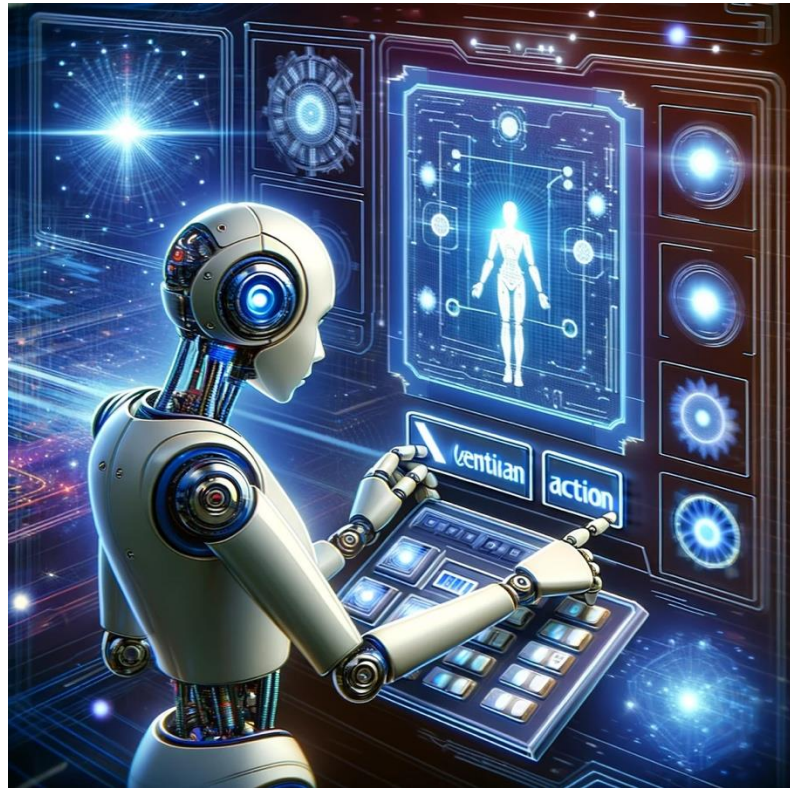
Trace Forecasting

Carbon Budget

Searched
Architecture

Hypervolume

Reinforcement Learning -- Action

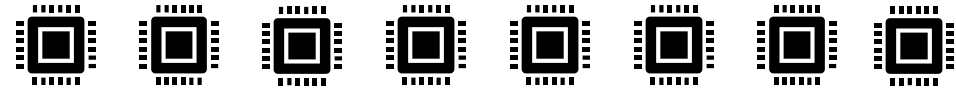


RL agent

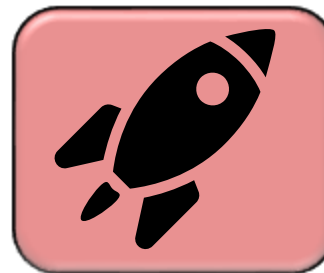


Ratio calculation

GPUs



High energy tasks



Low energy tasks



Reinforcement Learning -- Reward

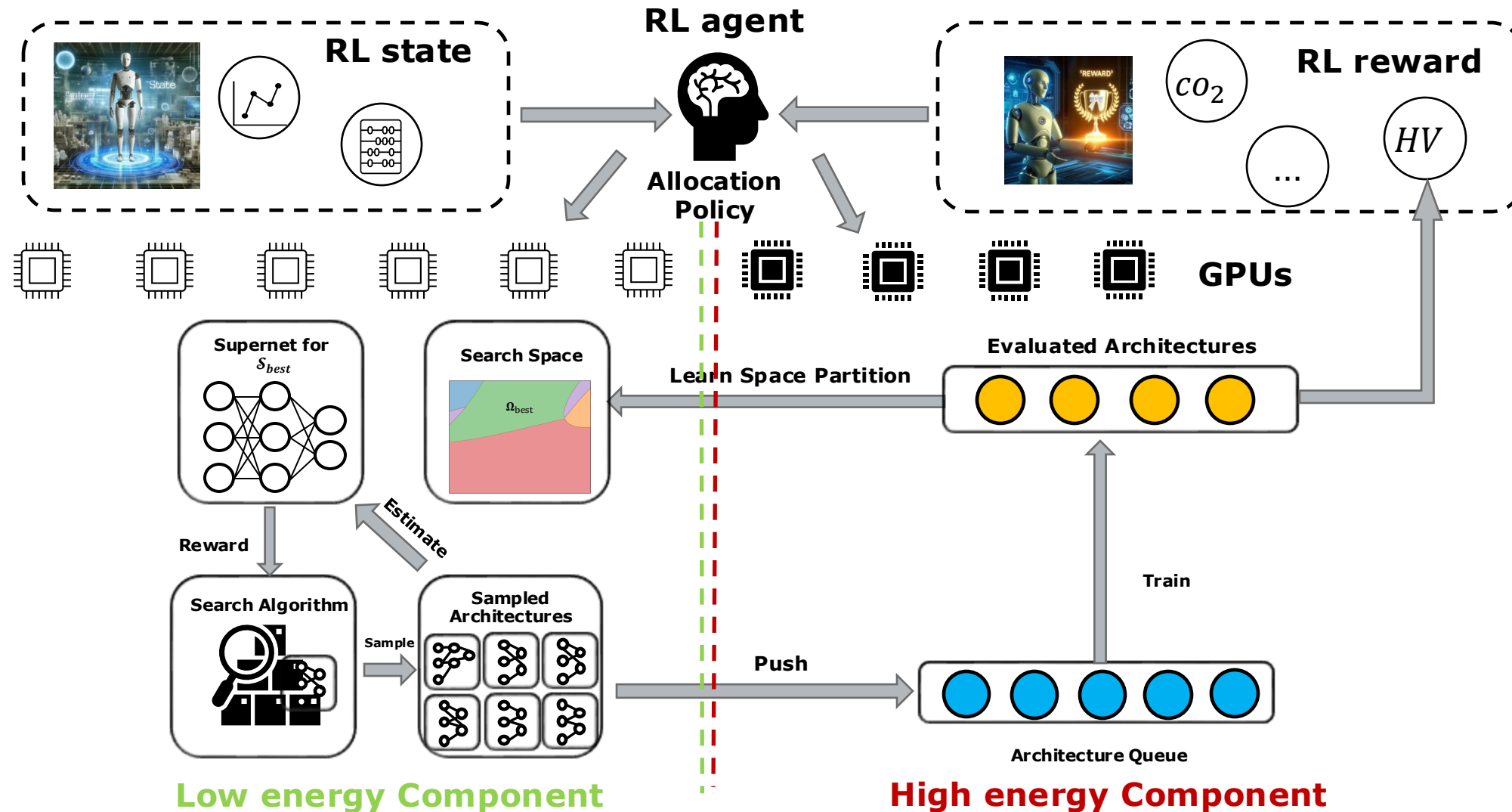


Carbon Cost

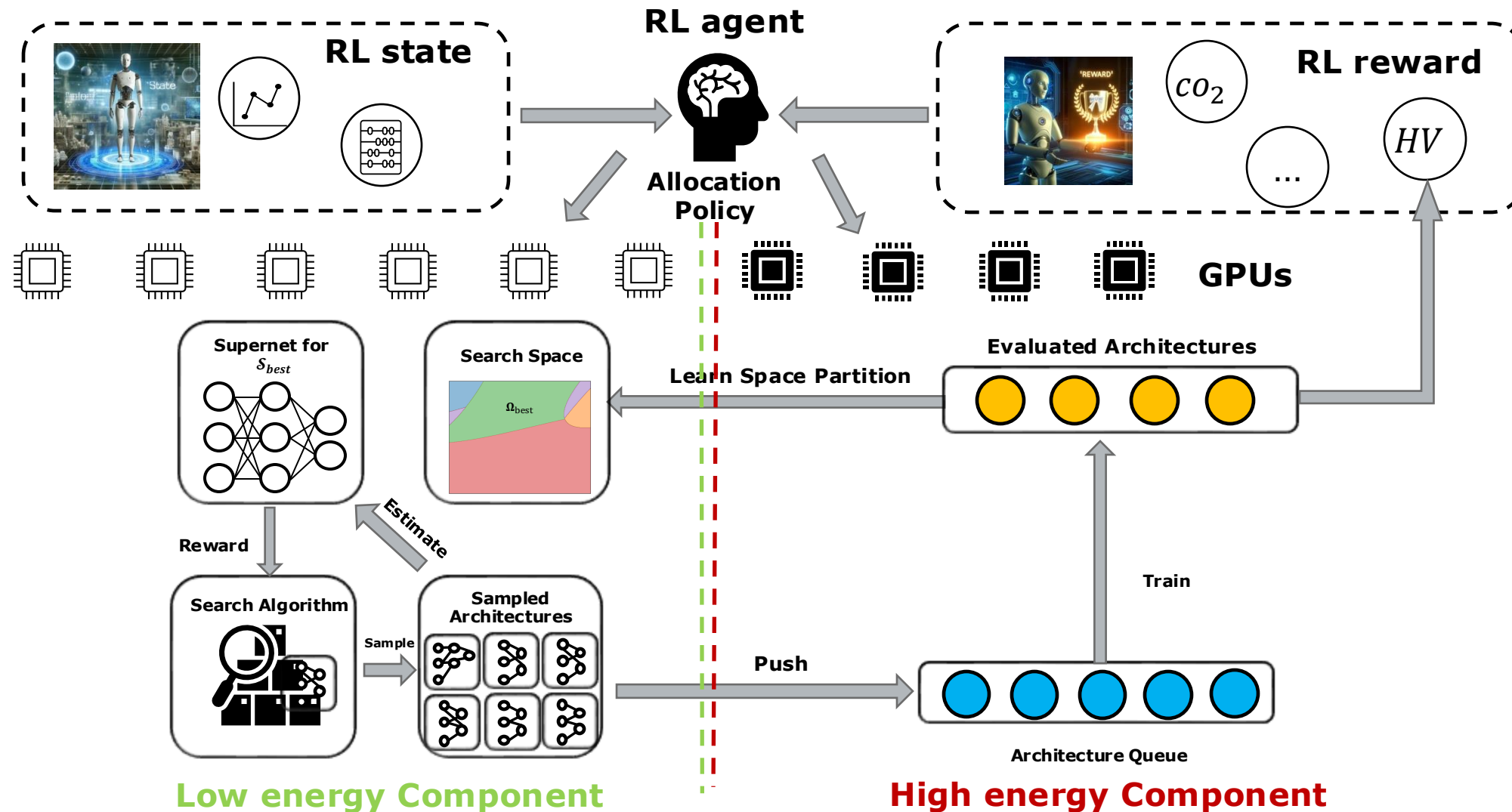
Hypervolume
Increase

#New Samples

How to allocate GPU resource



How to allocate GPU resource





THANK YOU