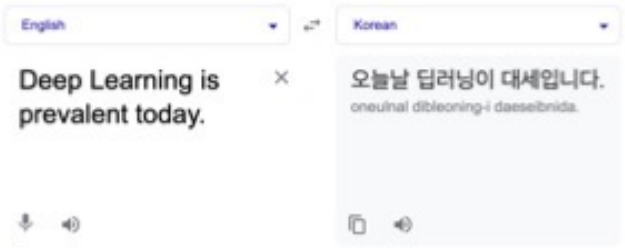
Zeus: Understanding and Optimizing GPU Energy Consumption of DNN Training

NSDI’23

2024.4

Background

- Deep learning applications is everywhere
  - Image processing
  - Autonomous driving
  - Machine translation
  - Image/text generation



TESLA

stability.ai

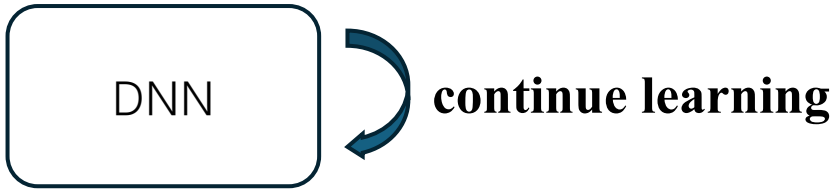
OpenAI

DNN Energy Consumption is Skyrocketing

- Training LLMs（GPT-3）

120 years of electricity for a household

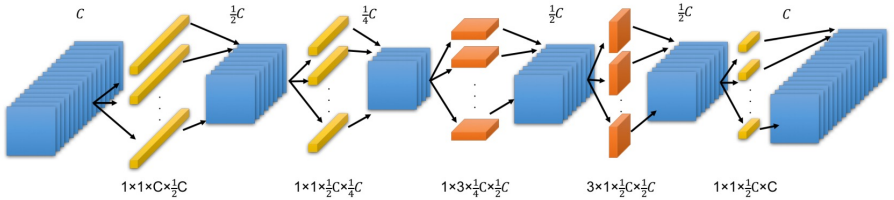
- Re-training demands



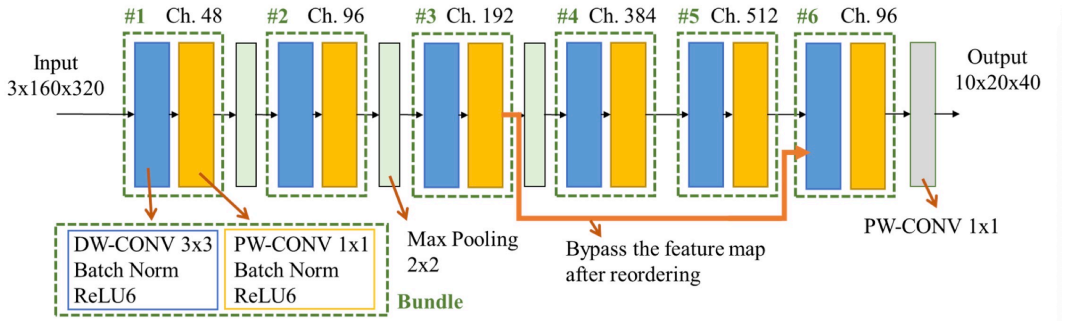
Re-training is commonplace(e.g. every hour)

Existing Energy-Efficient Solution

- New energy-efficient DNN architectures



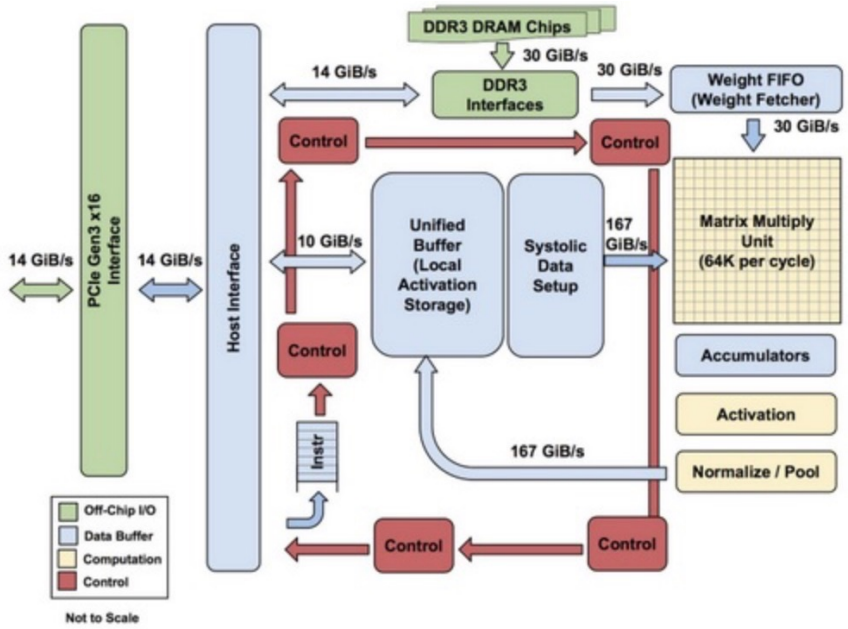
SqueezeNext



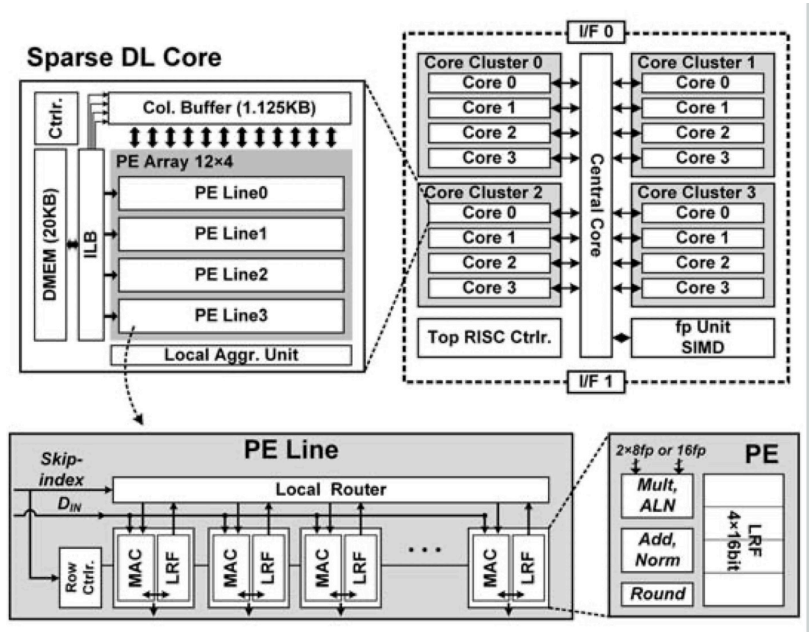
SkyNet

Existing Energy-Efficient Solution

- New energy-efficient HW architectures



TPU



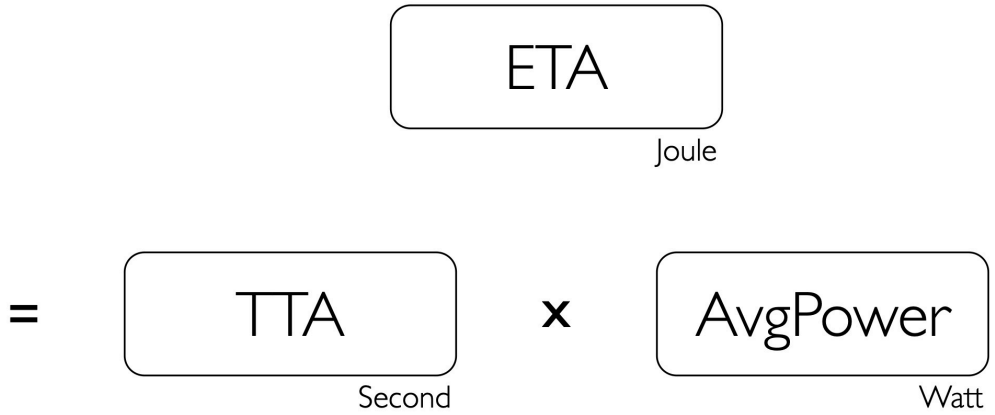
LNPU

Existing Energy-Efficient Solution

- Offline profiling and power model fitting.
- Confined to GPU power configuration knobs.

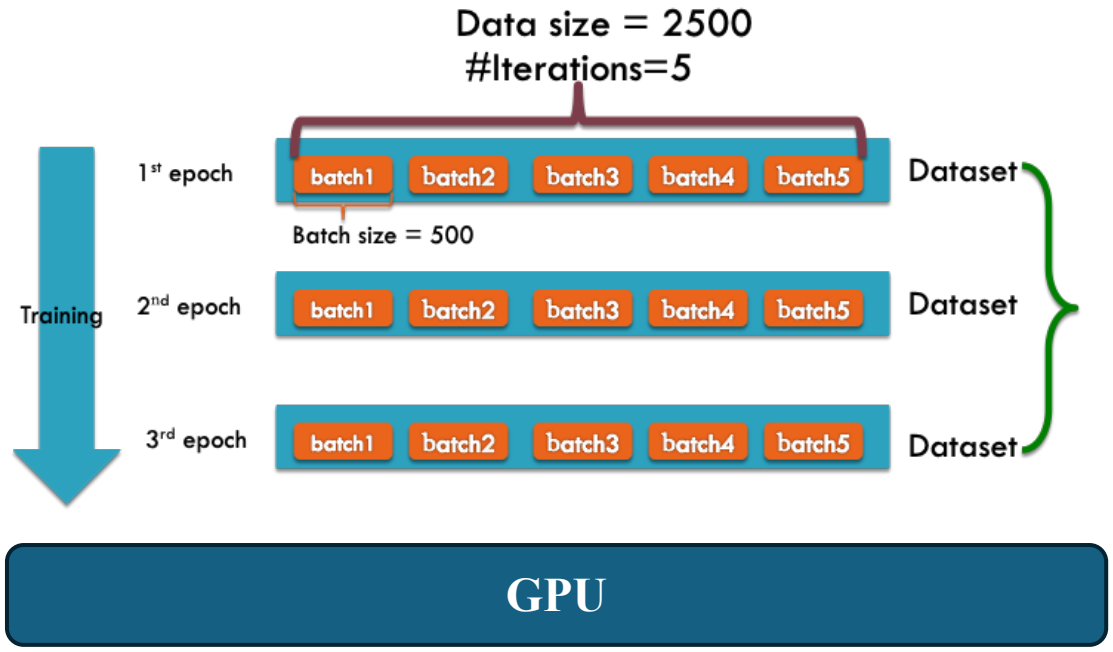
Consider the energy consumption of training

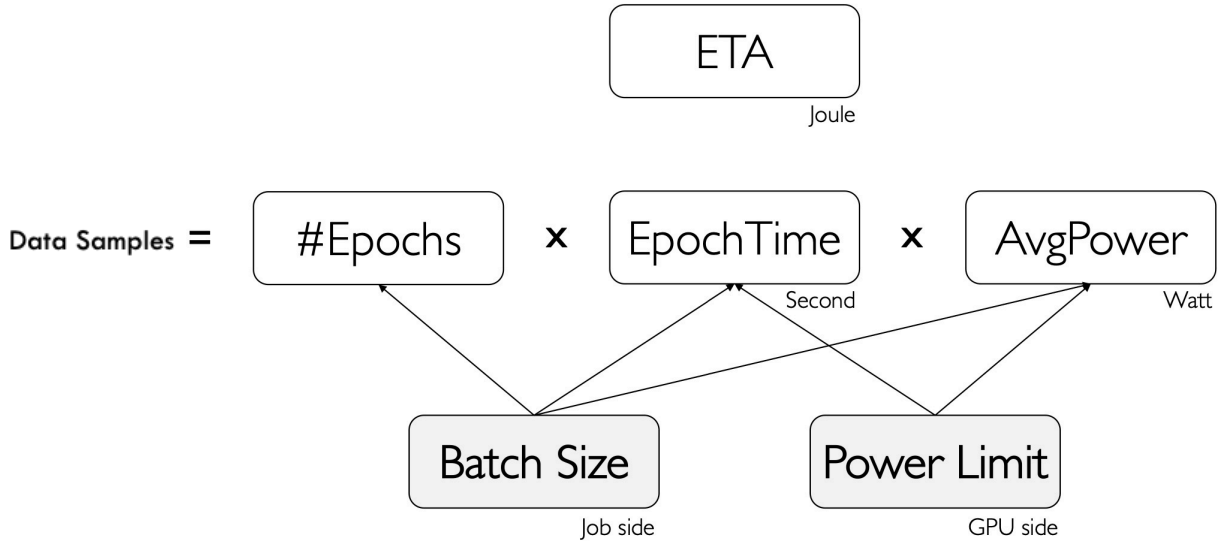
- Energy to Accuracy(ETA)
  - Energy needed to reach the user-specified **target accuracy**.
  - Energy-counterpart of **Time to Accuracy (TTA)**



Consider the energy consumption of training

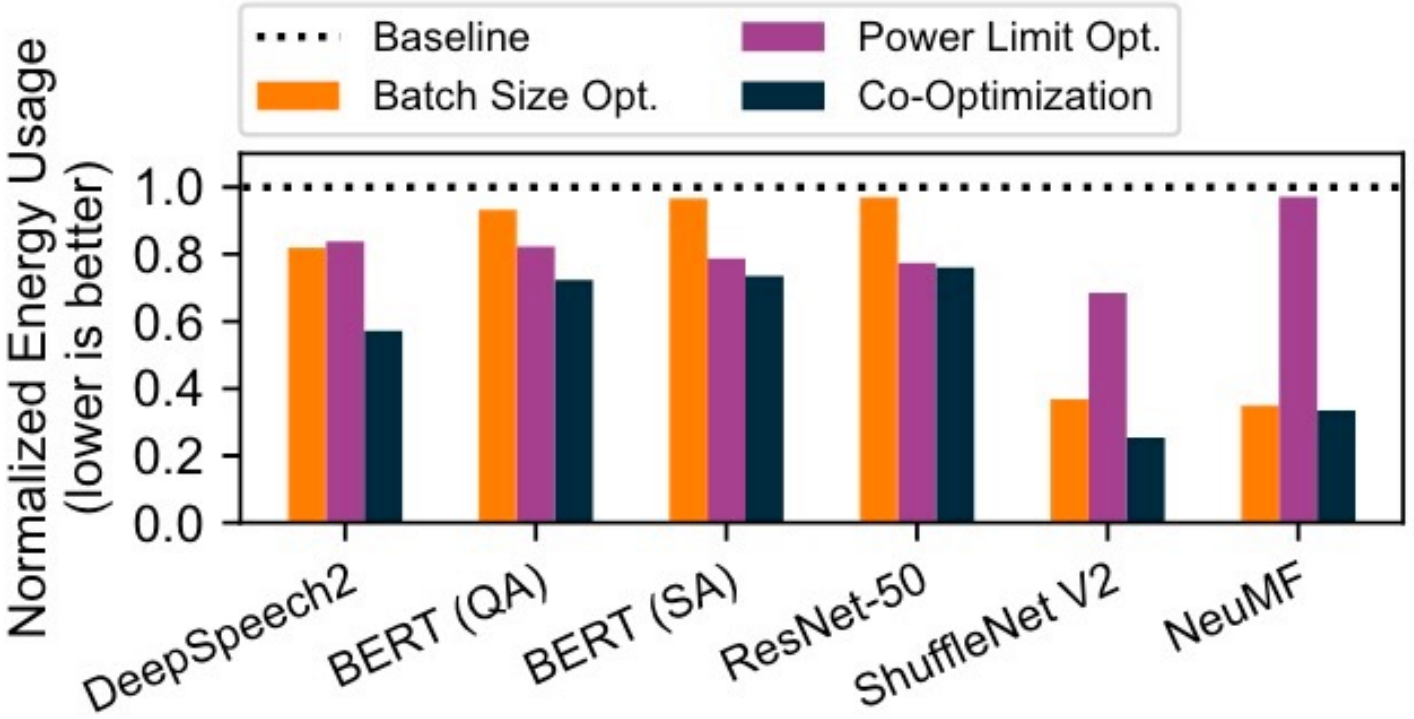
- Energy to Accuracy(ETA)
  - Energy needed to reach the user-specified **target accuracy**.
  - Energy-counterpart of **Time to Accuracy (TTA)**





Opportunities for Improving Energy Efficiency

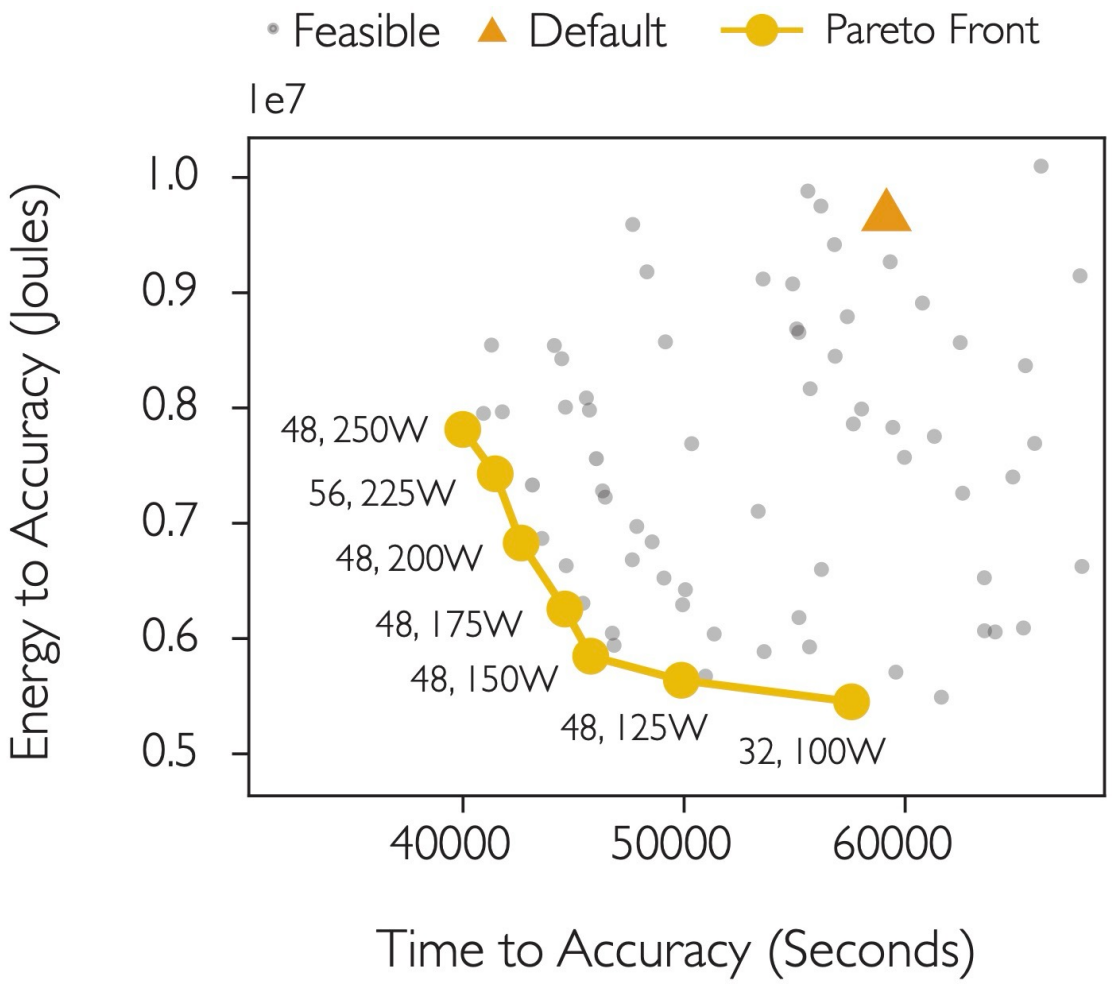
- Training terminates when the DNN reaches its original target accuracy.
  - **24~75% energy reduction with optimization**



Relationship between time and energy

- Efficient time and energy show a **trade-off** .
- Optimize metric

$$\text{Cost} = \eta \cdot \text{ETA} + (1 - \eta) \cdot \text{MaxPower} \cdot \text{TTA}$$

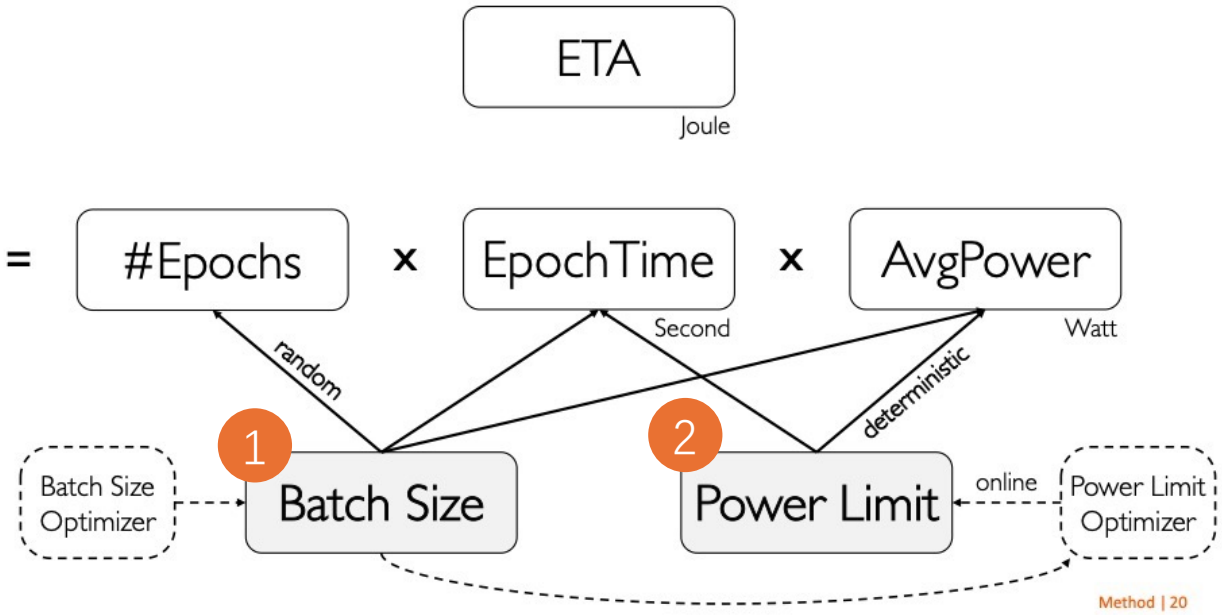


Decoupling Batch Size and Power Limit

$$\mathtt{ETA}(b,p) = \mathtt{TTA}(b,p) \times \mathtt{AvgPower}(b,p)$$

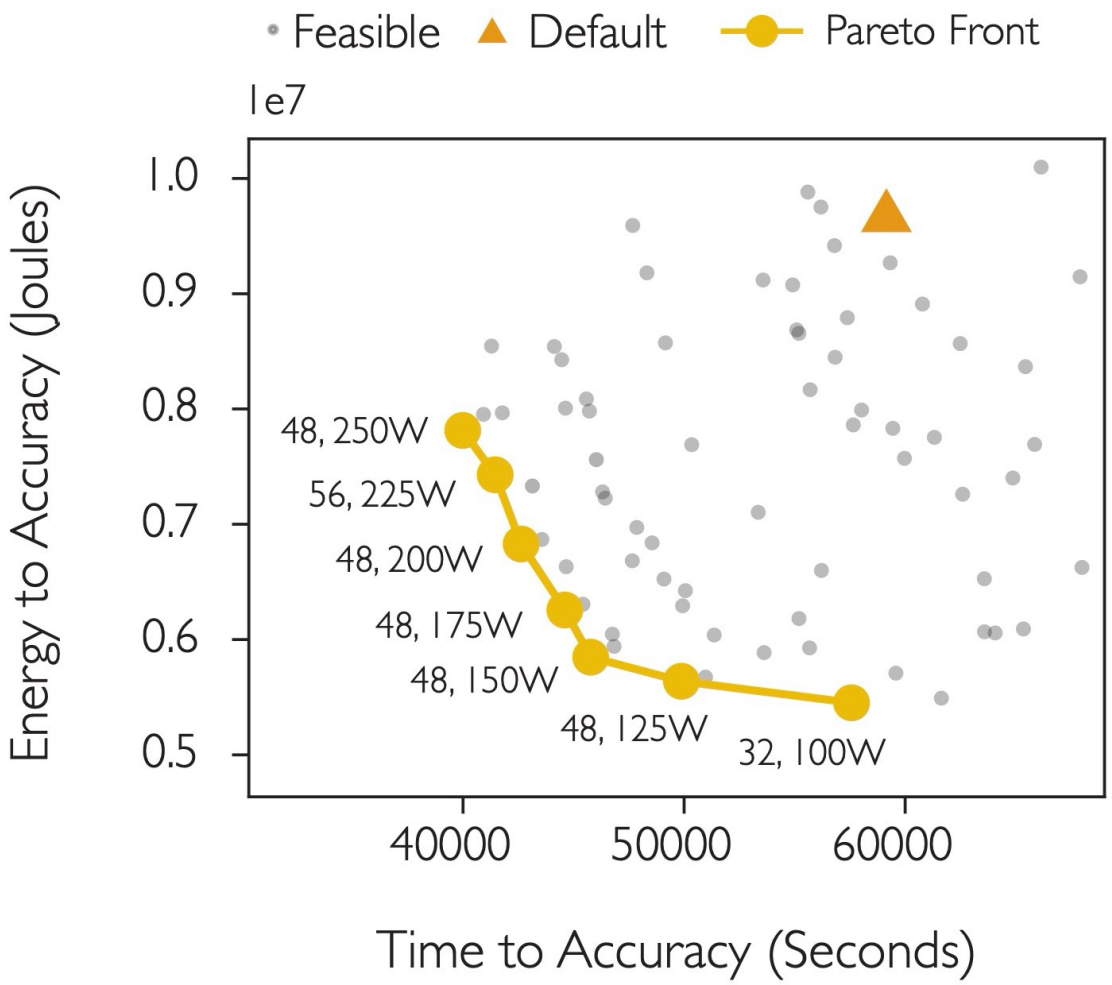
$$\mathtt{Cost} = \eta \cdot \mathtt{ETA} + (1-\eta) \cdot \mathtt{MaxPower} \cdot \mathtt{TTA}$$

$$\begin{aligned} &C(b,p;\eta) \\ &= (\eta \cdot \mathtt{AvgPower}(b,p) + (1-\eta) \cdot \mathtt{MAXPOWER}) \cdot \mathtt{TTA}(b,p) \\ &= \mathtt{Epochs}(b) \cdot \frac{\eta \cdot \mathtt{AvgPower}(b,p) + (1-\eta) \cdot \mathtt{MAXPOWER}}{\mathtt{Throughput}(b,p)} \end{aligned}$$

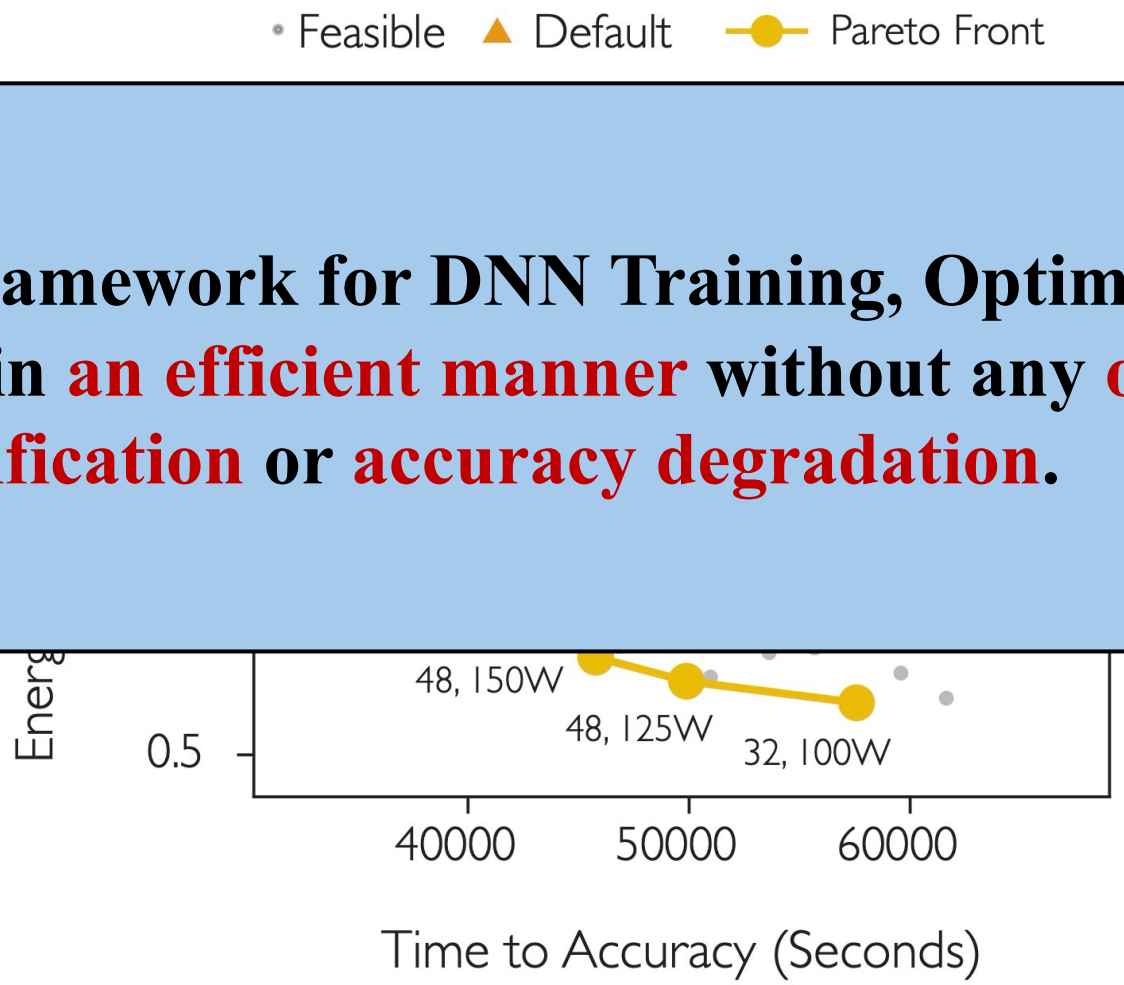


Relationship between time and energy

- Offline Profiling is **time-consuming**.
  - Different DNN model
  - Different Dataset
  - Different GPU



Relationship between time and energy



Build An Energy Optimization Framework for DNN Training, Optimize the cost of an arbitrary DNN model in an efficient manner without any offline profiling, hardware modification or accuracy degradation.

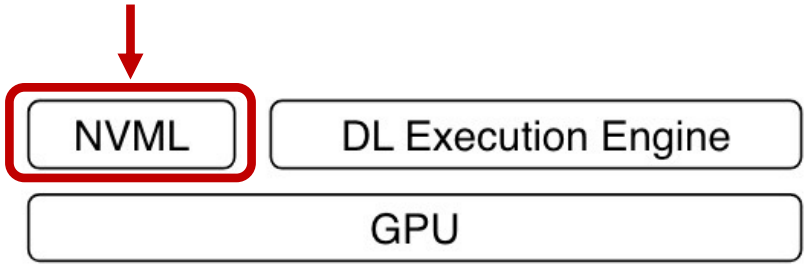
System Overview

- Architecture

System Overview

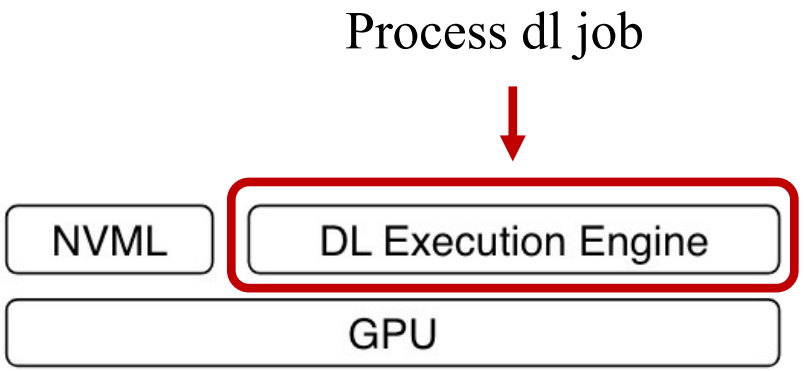
- Architecture

Monitor energy consumption



System Overview

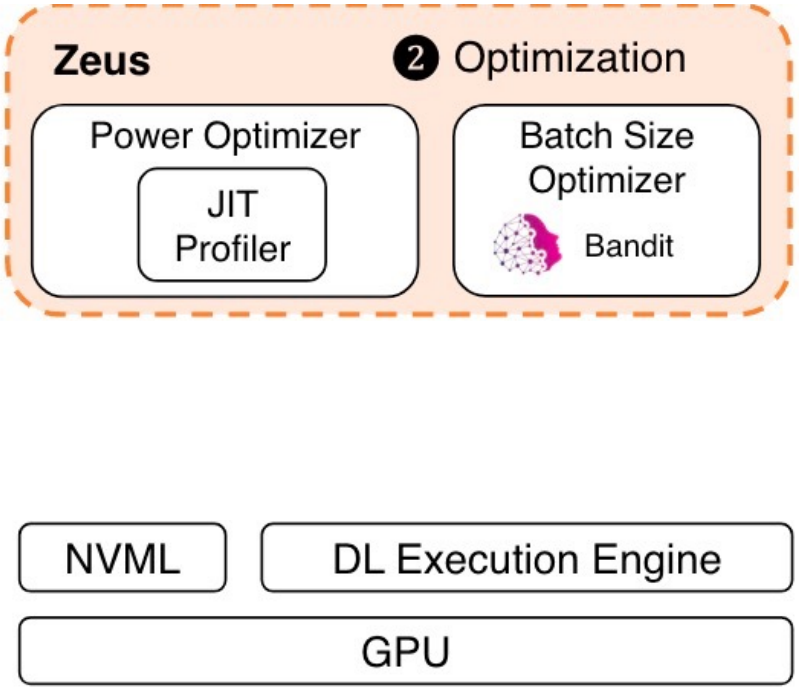
- Architecture



System Overview

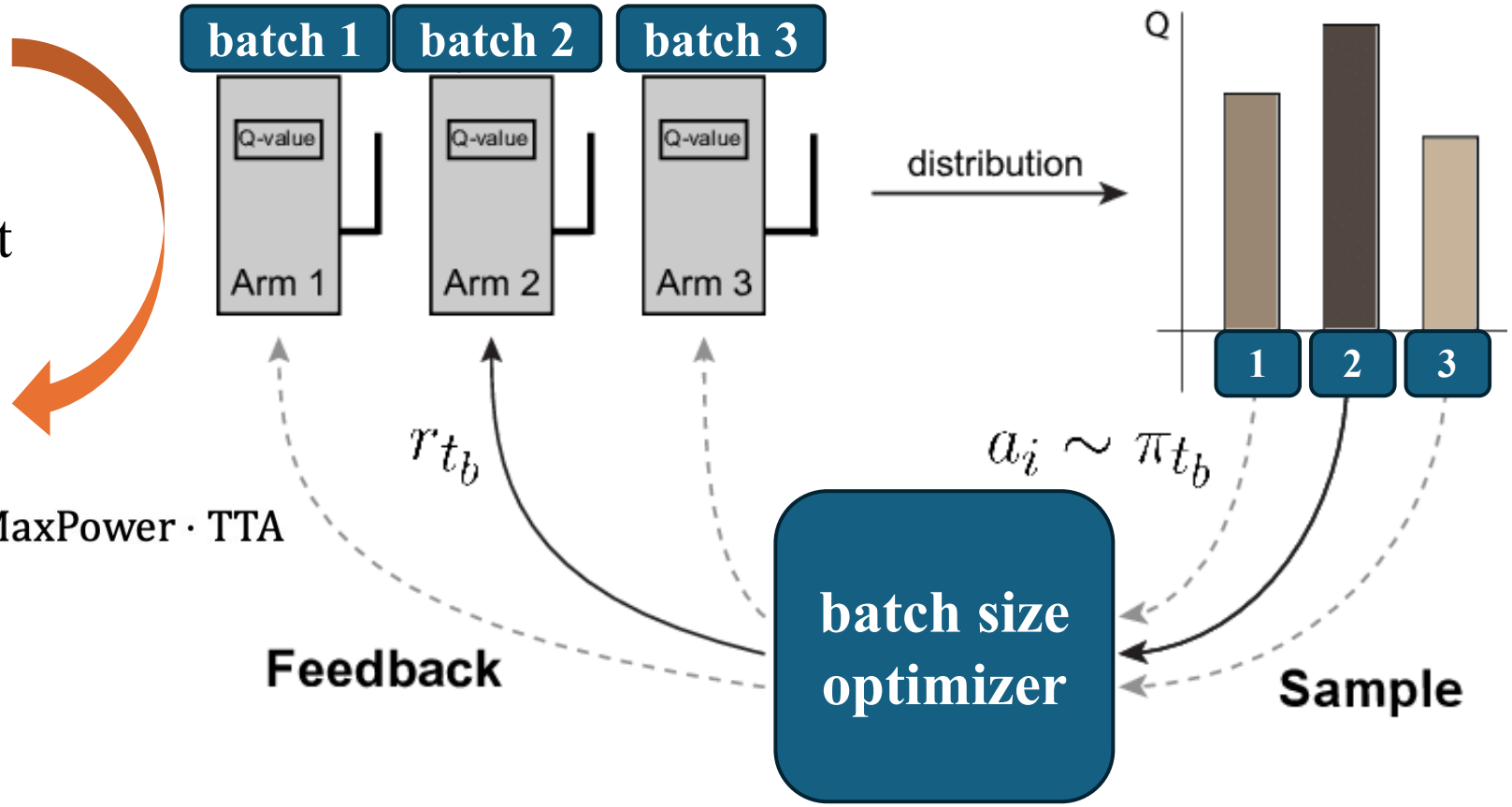
- Zeus Architecture
  - JIT Profiler
    - Profiles the **power** and **throughput** of each power limit
    - **Five seconds** per power limit is enough
  - Batch Size Optimizer
    - Multi-armed bandit

a **software-level** optimization framework.



Batch Size Optimizer

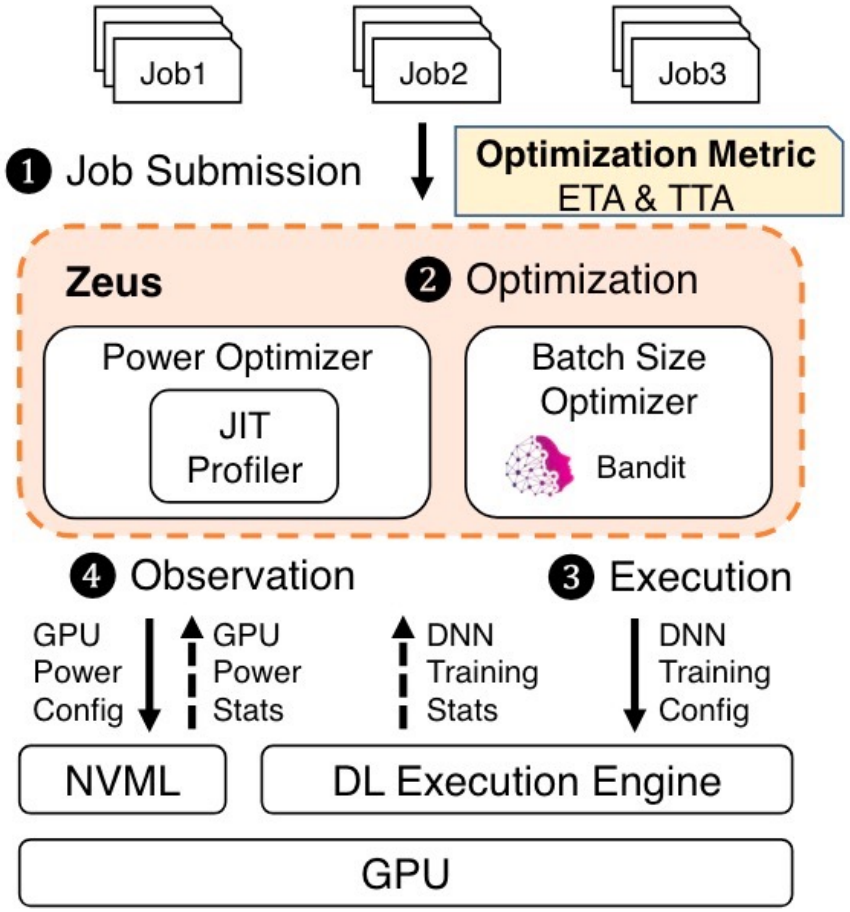
- Multi-armed bandit
  - Models **cost** as a Gaussian random variable
  - online, lightweight



$$\text{Cost} = \eta \cdot \text{ETA} + (1 - \eta) \cdot \text{MaxPower} \cdot \text{TTA}$$

System Overview

- Workflow of Zeus
  - Job submission
  - Optimization
  - Execution
  - Observation



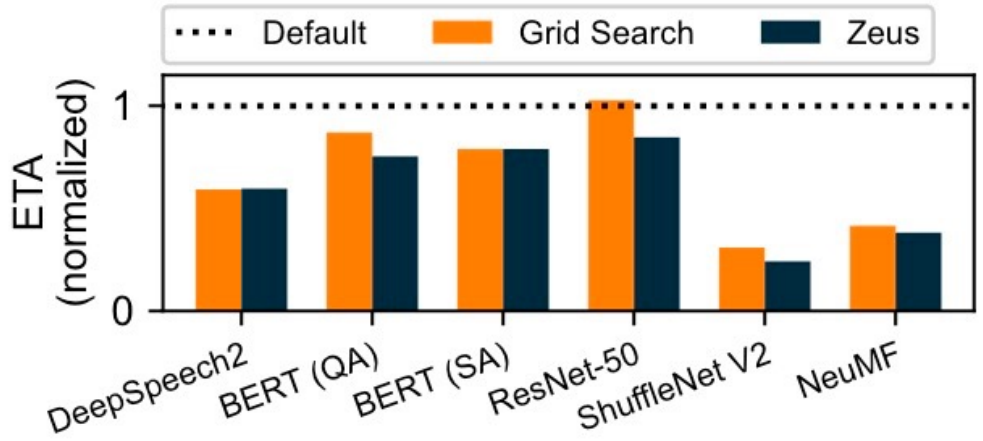
Experiment results

Task	Dataset	DNN
Speech Recognition	LibriSpeech	DeepSpeech2
Question Answering	SQuAD	BERT
Sentiment Analysis	Sentiment140	BERT
Image Classification	ImageNet	ResNet-50
Image Classification	CIFAR-100	ShuffleNet-v2
Recommendation	MovieLens-1M	NeuMF

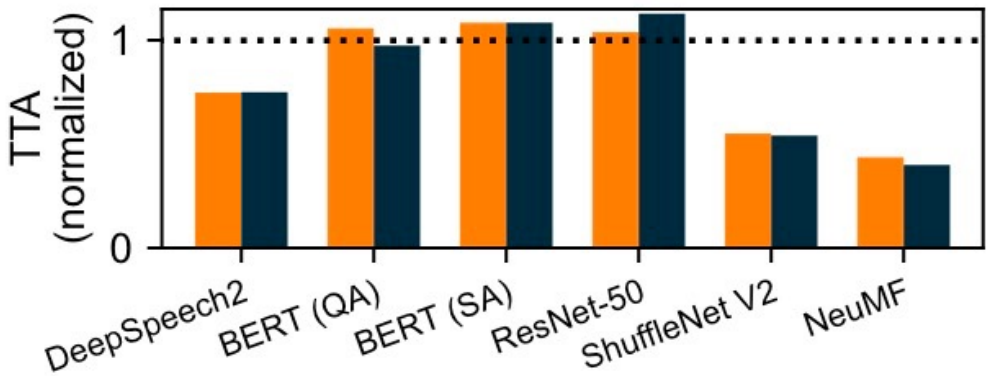
GPU	Arch
NVIDIA A40	Ampere
NVIDIA V100	Volta
NVIDIA RTX6000	Turing
NVIDIA P100	Pascal

Experiment results

- ETA(lower is better) & TTA(higher is better)



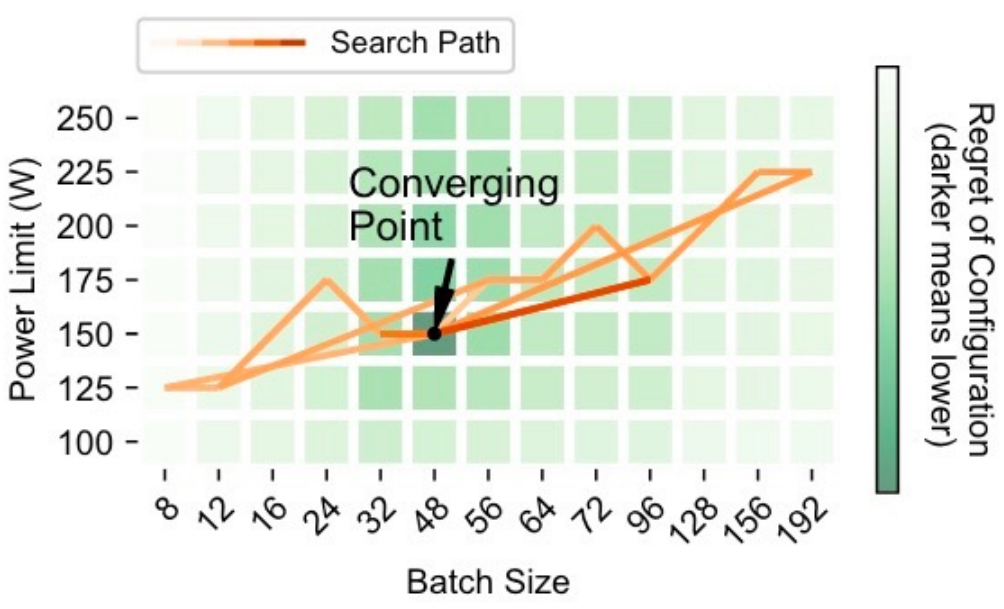
(a) *Energy Consumption*



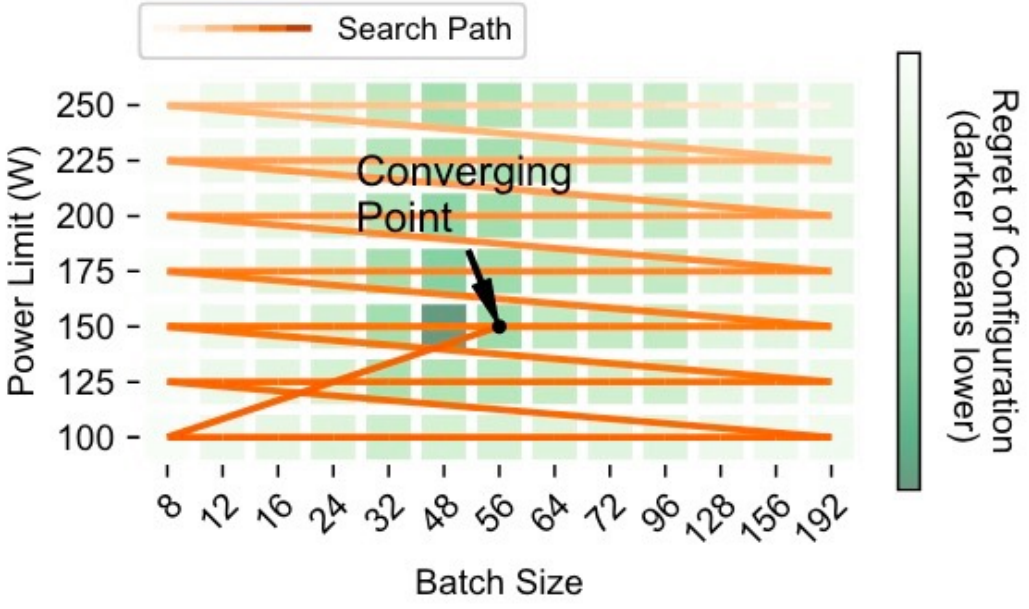
(b) *Training Time*

Experiment results

- Search paths



(a) *Zeus*



(b) *Grid Search*

Experiment results

- Handling Data Drift
  - Spikes in ETA and TTA (signaling that the current batch size may no longer be optimal) trigger the exploration of a batch size.