

CUT AI MODEL TRAINING COSTS IN **HALF**.

A NEXT-GENERATION AI TRAINING LOOP - NOT JUST ANOTHER OPTIMIZER.

Same model. Same data. Same framework. A different training loop.

2x

FEWER TRAINING STEPS

to the same model quality

~50%

LOWER TRAINING COST

to the same model quality

Measured: AERO at 1x Chinchilla = AdamW at 2x Chinchilla = 2.55 validation loss.

TWO ADVANTAGES THAT STACK

1

FEWER STEPS

2.0x token efficiency

Hardware-independent; measured at matched validation loss.

2

FASTER DISTRIBUTED STEPS

1.86x per-step throughput

Measured on 4 nodes / 32xA100, 45 GB/s inter-node link.

On that cluster: 3.7x end-to-end wall-clock speedup at matched validation loss.

FITS YOUR EXISTING TRAINING STACK

TRAINING WORKLOADS

Pretraining, post-training, RL, transformers and other architectures.

HARDWARE AGNOSTIC

Designed for existing compute; measured on H100 and A100.

YOUR ENVIRONMENT

Your model, data, framework and infrastructure stay in place.

THE SIMPLEST WAY TO UNDERSTAND AERO IS TO TEST IT.

Run AERO head-to-head in your existing environment.

Same setup. Decide with your own results.

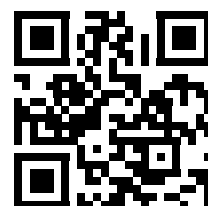


DevOpt
Labs

devoptlabs.com

REQUEST AN EVALUATION

Scan or visit
devoptlabs.com

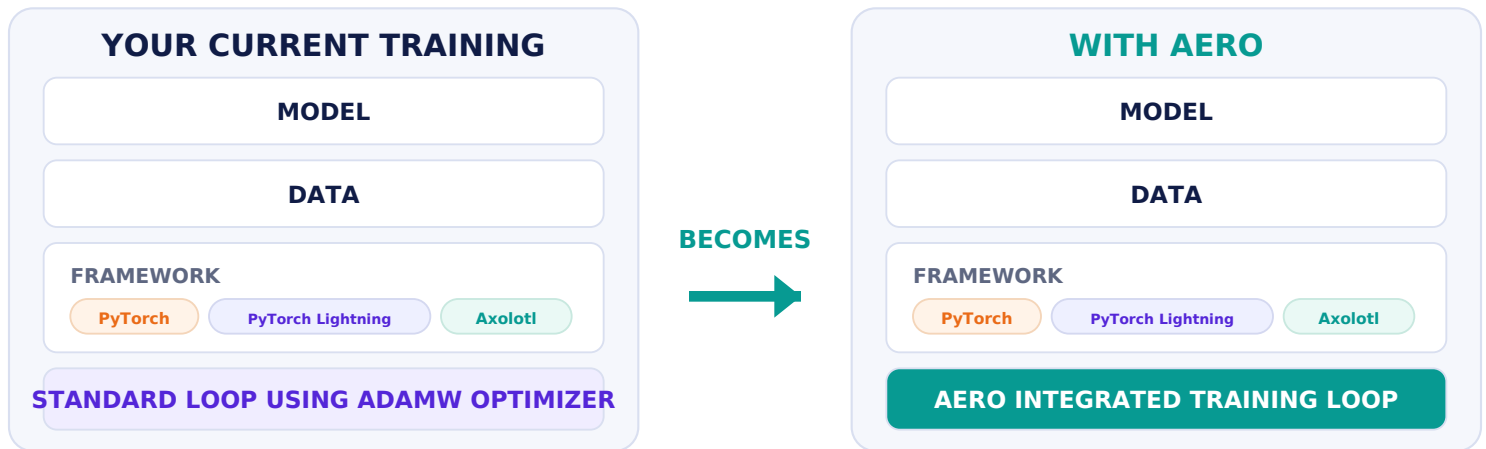


AERO FITS THE WAY YOU TRAIN TODAY



No new model architecture. No new dataset.
No wholesale rewrite of your training stack.

WHAT CHANGES - AND WHAT DOES NOT



CHANGE THE TRAINING LOOP. KEEP THE REST.

AERO replaces the standard forward / backward / optimizer step as one integrated training system.

1 PLUG IN

Use your existing model, data, framework and infrastructure.

Replace the standard training loop with AERO.

2 RUN HEAD-TO-HEAD

Use the setup you already trust: same model, data, framework, hardware and quality criterion.

No benchmark leap of faith.

3 DECIDE WITH RESULTS

Compare:

- time to target quality
- total training cost
- final model quality
- stability

Your models and data remain in your environment and under your control.

ACROSS TRAINING STAGES AND MODEL ARCHITECTURES

TRAINING STAGES

Pretraining | Continued pretraining | Fine-tuning / post-training | Reinforcement learning

ARCHITECTURES Transformers | Vision / CNNs | Diffusion | Other model architectures

MORE THAN AN OPTIMIZER SWAP

AERO integrates the training loop, reaches matched quality in fewer steps, and reduces optimizer / update traffic when distributed training is communication-bound.

THE SIMPLEST WAY TO UNDERSTAND AERO IS TO TEST IT.

Run AERO against your current training loop on your workload.



REQUEST AN EVALUATION

devoptlabs.com

