Your review report

Manuscript

Characterization of high-resolution AI data center training workloads on single and multiple GPU nodes

Feedback for the author(s)

Review file(s)

No files added.

Comments to the author(s)

This paper presents the first high-resolution (50 Hz) dataset characterizing AI training workloads across consumer-grade and datacenter-grade GPUs, enabling sub-second power dynamics analysis for grid connection impact assessment.

1. The paper uses HWiNFO (100 ms) for single-machine and Python scripts (20 ms) for node-scale experiments. How should researchers handle this 5× sampling rate difference when comparing or fusing data across platforms? Has any resampling or time alignment been performed?

2. Figure 5 shows "Utilization/Usage (%)" with values exceeding 100% (y-axis scale up to 150). What is the difference between "Utilization" and "Usage"? Why can values exceed 100%—is this multi-core summation or a different calculation?

3. Figure 8 labels "Node Power (W)", but the paper defines it as "sum of all individual GPU powers" and states CPU power could not be captured. Should this be renamed to "Total GPU Power" to avoid confusion with true node power (which includes CPU, memory, NVMe, etc.)?

4. Table 2 lists OS as "Ubuntu Server 22.04 arm64, x86-64" (confusing) and CPU as "Intel Xeon 208 vCPU" (suggesting virtualization). Please clarify the instance type and supplement software stack details (driver, CUDA, PyTorch versions) to support reproducibility?

5. Table 4 lists multiple hyperparameter values and "Parallelization Settings: ds(Z1, Z2, Z3)", are these full grid combinations or sampled subsets? Which session corresponds to which parameter set? What do Z1/Z2/Z3 represent?

6. The paper states node-scale cannot capture CPU power/temp, but Listing 1 still outputs these fields; are they empty or not provided in node data?

7. An isolated "aa" string appears after Figure 9 (around line 595).

Confidential feedback for the Editor

Your recommendation

- ***Revise***

Confidential comments to the Editor

High-resolution (50 Hz) AI workload power data across RTX 3060 and H100/B200 GPUs; 72 sessions, 1.8M+ samples; valuable for grid impact assessment. Main Issues: Metric definitions inconsistent ("Node Power" = GPU only; values >100%); reproducibility explanation is insufficient; code listing errors prevent direct reuse.