

PR #2300 完整报告

radixark/miles

scripts: enable the Miles dashboard in the quick-start launcher

合并时间: 2026-08-11 02:27

原文链接: <http://prhub.com.cn/radixark/miles/pull/2300>

执行摘要

- 一句话: Quick Start 启动脚本默认开启 Miles dashboard 遥测
- 推荐动作: 值得快速浏览, 但无需精读。该 PR 展示了如何在启动脚本中按参数组形式接入新特性, 并遵循已有的 `<output>/dump_details` 目录约定。若关注 shell 启动脚本的可测试性, 可留意该脚本未被现有快照测试覆盖, 后续可考虑补上。

功能与动机

PR body 说明这是新用户第一个运行的脚本, 希望开箱即用地获得 dashboard 的 GPU timeline 与 token 级轨迹视图; 文档 Quick Start 页面 (#2298) 也直接指向该 dump 路径供 `miles.dashboard.serve` 渲染。作者特意不加 `--use-rollout-entropy`, 避免增加计算开销, 仅在有该标志时缺失 per-token entropy 的日志警告。

实现拆解

1. 在 `scripts/run-qwen3-4B.sh` 的参数组区域 (PROMETHEUS_ARGS 之后、MISC_ARGS 之前) 新增 DASHBOARD_ARGS 数组, 包含 `--use-miles-dashboard` 和 `--dump-details /root/Qwen3-4B_miles/dump_details`。
2. 在文件末尾的 `ray job submit` 命令行中, 将 `${DASHBOARD_ARGS[@]}` 插入到 `${PROMETHEUS_ARGS[@]}` 与 `${MISC_ARGS[@]}` 之间, 确保参数按组顺序传递给 `train.py`。
3. `--dump-details` 路径放在 `recipe` 的 checkpoint 目录 `/root/Qwen3-4B_miles/` 下, 遵循 Python launcher 的 `<output>/.../dump_details` 约定。
4. 无配置、测试或部署配套改动; PR body 中的测试计划为手动的 `bash -n`、运行验证与 `dashboard.serve` 渲染检查。

关键文件:

- `scripts/run-qwen3-4B.sh` (模块 启动脚本; 类别 other; 类型 core-logic; 符号 DASHBOARD_ARGS, `--use-miles-dashboard`, `--dump-details`): quick-start 启动器是用户首次运行的入口, 本次在此加入 dashboard 遥测参数, 使遥测默认开启。

关键符号: 未识别

关键源码片段

scripts/run-qwen3-4B.sh

quick-start 启动器是用户首次运行的入口，本次在此加入 dashboard 遥测参数，使遥测默认开启。

```
# 位于 scripts/run-qwen3-4B.sh 的参数组区域
# Prometheus 监控参数（默认注释掉，用户可按需开启）
PROMETHEUS_ARGS=(
  # --prometheus-run-name "qwen3-4b-exp"
)

# 新增：Miles dashboard 参数组
# --use-miles-dashboard 开启 dashboard 遥测（fire-and-forget，毫秒级开销）
# --dump-details 指定 JSONL 输出目录，位于 recipe 的 checkpoint 根目录下
DASHBOARD_ARGS=(
  --use-miles-dashboard
  --dump-details /root/Qwen3-4B_miles/dump_details
)

MISC_ARGS=(
  # default dropout in megatron is 0.1
  --attention-dropout 0.0
)

# 在 ray job submit 命令末尾按顺序展开各参数组
# DASHBOARD_ARGS 插在 PROMETHEUS_ARGS 与 MISC_ARGS 之间，确保透传给 train.py
ray job submit --address="http://127.0.0.1:8265" \
  "${EVAL_ARGS[@]}" \
  "${SGLANG_ARGS[@]}" \
  "${PROMETHEUS_ARGS[@]}" \
  "${DASHBOARD_ARGS[@]}" \
  "${MISC_ARGS[@]}"
```

评论区精华

无 review 评论；Zhichenzzz 直接 APPROVED，未提出任何问题或建议。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低。变更仅新增两个参数并传递到 train.py，不涉及核心逻辑。潜在风险点：
 - 1) --dump-details 路径为硬编码 /root/Qwen3-4B_miles/，若用户改变 recipe 输出目录，dump 数据会写到默认位置，与 checkpoint 目录不一致；
 - 2) scripts/run-qwen3-4B.sh 是 shell 脚本，本次改动没有对应的快照测试（同仓历史 PR 已建立 launch_scripts 快照体系，但该脚本不在覆盖范围内）；
 - 3) dashboard 遥测虽为 fire-and-forget，若 --use-miles-dashboard 在异常环境下导致额外日志或依赖缺失，可能影响首次运行体验，但概率很低。

- 影响：影响范围为 quick-start 用户：新用户运行 run-qwen3-4B.sh 后会自动在 /root/Qwen3-4B_miles/dump_details/ 下生成遥测 JSONL，可通过 miles.dashboard.serve 在 http://localhost:7788 查看。对现有用户无破坏性影响，仅新增了两行参数。对团队而言，这一步完成了 Quick Start 到 dashboard 的链路打通，为 #2298 文档提供可运行的基础。影响程度低。
- 风险标记：配置变更无自动化测试覆盖，硬编码 dump 路径

关联脉络

- PR #2298 docs: Quick Start page: PR body 明确提到 #2298 是叠在本 PR 之上的文档页面，指向 miles.dashboard.serve --dump-details /root/Qwen3-4B_miles/dump_details，本 PR 为它提供可运行的数据来源。