

PR #1919 完整报告

radixark/miles

examples: add swe-agent-harbor-daytona (Harbor sandboxes on Daytona)

合并时间: 2026-08-07 05:53

原文链接: <http://prhub.com.cn/radixark/miles/pull/1919>

执行摘要

- 一句话: 新增 Daytona 云沙箱 SWE-Agent 示例到 experimental
- 推荐动作: 值得精读。重点看两点: 一是 '何时应该放弃自己的启动器而复用兄弟示例' 的决策过程——run.py 因单一参数被移除而整体删除, 是一个避免重复维护的好案例; 二是 README 中沉淀的实操知识 (per-turn 响应上限、raw_reward 与 rewards 的区别、trial 目录 mtime 陷阱) 对任何跑 agentic RL 长任务的人都有直接价值。

功能与动机

PR body 明确说明: Daytona 沙箱方案是 "the practical option when the trainer runs on a GPU node where you cannot run Docker-in-Docker", 即训练节点无法运行 Docker 时, 任务评测沙箱放到 Daytona 云端, agent-server 主机只需出站 HTTPS。另一个动机是 terminus-2 agent 以宿主进程运行并自行调用模型端点, 需要追加 user 与 tool 轮次; 该 flag 曾是需要独立启动器的原因, 但 PR #1818 已移除该参数 (并新增测试断言其被拒绝), 这直接推动最终删除 run.py。

实现拆解

1. 新增示例目录与 README: examples/experimental/swe-agent-harbor-daytona/README.md 是变更入口, 说明与 examples/swe-agent 的两点差异 (Daytona 云沙箱、terminus-2 宿主 agent), 并强调训练侧完全复用: 数据准备用 examples/swe-agent/download_and_process_data.py --agent-name terminus-2, 训练启动用 examples/swe-agent/run.py 不变, Daytona 只由 agent server 的环境变量决定。
2. 启动脚本 launch_agent_server.sh: 在 harbor-framework/harbor 的 harbor-miles-v0.20.0 分支下启动 agent server, 设置 HARBOR_ENV_TYPE=daytona、HARBOR_DAYTONA_AUTO_SNAPSHOT=1 (首用后快照镜像)、OPENAI_API_BASE 指向本机 30000 端口 session server、AGENT_MAX_INPUT_TOKENS / AGENT_MAX_OUTPUT_TOKENS / HARBOR_RESPONSE_LENGTH_POLICY=abort 等与训练侧 --rollout-max-response-len、--max-seq-len 对齐的配置, --agent-timeout 默认 5400。
3. 索引更新: examples/README.md 追加一行指向新示例。
4. 演进与删除 un.py: 初版自带约 284 行的 un.py, 用于暴露 rollout-max-response-len、--max-tokens-per-gpu 等 tuning 参数并传 --tito-allowed-append-roles; review 中 nblintao 指出该参数已被 PR #1818 移除后, 作者确认其余配置均与

examples/swe-agent/run.py 硬编码一致，最终删除 run.py，行为完全等价。目录也从 examples/swe-agent-harbor-daytona 移入 examples/experimental/。

5. 测试配套：无自动化测试（README 声明 docs-and-example change）；以 py_compile、bash -n、ruff 检查，并由一次 8xH200 多日 GLM-4.7-Flash 端到端运行验证非零 rollout/raw_reward。

关键文件：

- examples/experimental/swe-agent-harbor-daytona/README.md（模块 示例文档；类别 docs；类型 documentation）：示例主文档，交代 Daytona/terminus-2 背景、完整操作步骤与多日运行踩坑总结，是 PR 价值的主要载体
- examples/experimental/swe-agent-harbor-daytona/launch_agent_server.sh（模块 启动脚本；类别 other；类型 core-logic）：唯一新增的代码逻辑文件，承载 Daytona 模式的全部环境开关与 agent server 启动参数
- examples/README.md（模块 示例索引；类别 docs；类型 documentation）：示例总索引追加新条目，帮助用户发现该示例

关键符号：未识别

关键源码片段

examples/experimental/swe-agent-harbor-daytona/launch_agent_server.sh

唯一新增的代码逻辑文件，承载 Daytona 模式的全部环境开关与 agent server 启动参数

```
#!/bin/bash
# 启动 Harbor agent server 的 Daytona 模式，需在 harbor-framework/harbor 的
# harbor-miles-v0.20.0 分支下运行，之后再用 examples/swe-agent/run.py 启动训练。
# trial 在 Daytona 云沙箱内评测，本机只需出站 HTTPS，无需 Docker daemon。
set -euo pipefail

# 必填环境变量：缺少时直接报错退出
: "${DAYTONA_API_KEY:?set DAYTONA_API_KEY to a Daytona API key}"
: "${HARBOR_TASKS_DIR:?set HARBOR_TASKS_DIR to the directory holding Harbor task dirs}"

# 可选配置默认值
TRIALS_DIR="${TRIALS_DIR:-/tmp/harbor_trials}"
PORT="${PORT:-11000}"
MAX_CONCURRENT="${MAX_CONCURRENT:-32}"

# 核心开关：声明使用 Daytona 云沙箱
# HARBOR_DAYTONA_DISK_GB 需结合 Daytona 账号总磁盘配额设置
export HARBOR_ENV_TYPE=daytona
export HARBOR_DAYTONA_DISK_GB="${HARBOR_DAYTONA_DISK_GB:-10}"
# 首次使用后给任务镜像打快照，后续 trial 跳过镜像构建
# 该项能显著降低重复启动的等待时间
# 注意：快照会占用账号磁盘配额，并发沙箱数乘磁盘大小需低于配额
export HARBOR_DAYTONA_AUTO_SNAPSHOT=1

# terminus-2 以宿主进程运行并自行调用模型端点，
```

```
# 因此模型端点必须从本机可达，而不是从沙箱内可达
# 这里默认指向本机 30000 端口的 Miles session server
export OPENAI_API_KEY="{OPENAI_API_KEY:-dummy}"
export OPENAI_API_BASE="{OPENAI_API_BASE:-http://127.0.0.1:30000/v1}"
export OPENAI_BASE_URL="{OPENAI_API_BASE}"

# 以下两值与训练侧 --rollout-max-response-len、--max-seq-len 保持一致，
# 它们限制的是单轮输出而非整条轨迹，详见 README 的调节说明
# 注意：HARBOR_RESPONSE_LENGTH_POLICY=abort 下超限会导致整个 trial 得 0 分
export AGENT_MAX_INPUT_TOKENS="{AGENT_MAX_INPUT_TOKENS:-32768}"
export AGENT_MAX_OUTPUT_TOKENS="{AGENT_MAX_OUTPUT_TOKENS:-8192}"
export HARBOR_RESPONSE_LENGTH_POLICY=abort

mkdir -p "$TRIALS_DIR"

# 注意：--agent-timeout 5400 秒，而 examples/swe-agent/swe_agent_function.py
# 中硬编码 3600 秒，二者不一致已被 PR #2228 跟进为可配置
exec python miles_agent_server.py \
--host 0.0.0.0 \
--port "$PORT" \
--max-concurrent "$MAX_CONCURRENT" \
--agent-timeout 5400 \
--trials-dir "$TRIALS_DIR"
```

评论区精华

核心交锋集中在三点：一是 run.py 是否还值得存在——nblintao 直接以 PR #1818 为证质疑，Shi-Dong 确认后删除，这是本 PR 最重要的设计收敛；二是 nblintao 提议将 **swe-agent** 重命名为 **swe-agent-harbor-docker** 以区分后端，Shi-Dong 当场开 PR #2233 跟进；三是 nblintao 指出 launch_agent_server.sh 的 **--agent-timeout 5400** 与 **examples/swe-agent/swe_agent_function.py** 里硬编码的 3600 秒超时不一致，agent function 放弃后 server 仍在运行，Shi-Dong 开 PR #2228 跟进。另有关于 shell 脚本是否应收敛为 Python 的讨论，结论是脚本用于启动 agent server 而非训练任务，转 .py 主要为 CI 考虑，暂不处理。

- run.py 是否还需要存在（PR #1818 已移除 --tito-allowed-append-roles）（design）：Shi-Dong 确认后删除 run.py，复用 examples/swe-agent/run.py，其余配置默认值完全一致，行为等价
- 是否将 swe-agent 重命名为 swe-agent-harbor-docker（design）：Shi-Dong 采纳建议，开 PR #2233 处理重命名
- agent server 超时 5400 与 agent function 硬编码 3600 不一致（correctness）：Shi-Dong 开 PR #2228 将 agent-server trial 超时改为可配置
- shell 脚本是否应收敛为 Python（style）：保留 shell 脚本，未在此 PR 内转换，转为后续讨论话题
- README 规范：harbor 分支、相对路径、硬件说明（documentation）：全部落实：README 与脚本中改用仓库根路径、补充单节点 H200 说明、删除冗余排查行

风险与影响

- 风险:

1. 示例与文档无自动化测试覆盖: 变更纯示例性质, README 也明说没有现有测试覆盖, 后续 `examples/swe-agent/run.py` 或 Harbor 分支接口变化时, 本示例容易静默失配。
2. 依赖外部 harbor 分支: `harbor-miles-v0.20.0` 是 `harbor-framework/harbor` 的特定积分分支, 其 API 或环境变量变化会破坏启动脚本, 风险集中在 `launch_agent_server.sh`。
3. agent 超时配置不一致: `swe_agent_function.py` 硬编码 3600 秒, 而启动脚本给 `server` 设置 5400 秒, 二者不匹配会导致 agent function 提前放弃而 `server` 侧 trial 变为孤儿任务; 该问题已由 PR #2228 跟进修复。
4. 运维陷阱已有文档但依赖执行纪律: README 强调 per-turn 响应上限、`raw_reward` 判读、`trial` 目录 `mtime` 陷阱, 一旦读者跳过这些说明, 多日运行仍会误判进度或产出 0 分 trial。
 - 影响: 对用户: 提供了一条在无法运行 Docker-in-Docker 的 GPU 节点上跑 SWE-Agent/Harbor 训练的可执行路线, 并把多日调参踩坑总结成可复用的运维手册。对团队: review 直接催生了 PR #2233 (示例命名) 和 PR #2228 (超时可配置) 两个后续改进; 也确认了 '训练侧零改动、沙箱后端完全由 agent server 环境决定' 的架构抽象。对系统: 无运行时影响, 不触碰库代码, 风险面限定在示例目录。
 - 风险标记: 示例与文档无自动化测试覆盖, 依赖外部 harbor 分支 `harbor-miles-v0.20.0`, agent 超时配置不一致 (3600 vs 5400), 依赖 PR #1818 移除参数后的行为

关联脉络

- PR #2233 examples: rename swe-agent to swe-agent-harbor-docker: 由本 PR review 中 nblintao 的重命名建议直接产生, Shi-Dong 在讨论中明确 'Opened PR 2233 to handle the renaming'
- PR #2228 examples: make the agent-server trial timeout configurable: 由本 PR review 中发现的 5400 与 3600 超时不匹配问题引出, Shi-Dong 表示 'Opened a follow-up PR 2228 to address this comment'
- PR #1913 E2B sandbox backend (E2B Cloud / self-hosted AgentENV) + dedicated AgentENV recipe: 同为 agent 沙箱后端扩展, 与本 PR 一起印证 '沙箱后端由 agent server 环境决定、训练侧零改动' 的抽象方向