

PR #31757 完整报告

sgl-project/sglang

[AMD] Build Miles nightly ROCm images with docker/build.py and test ROCm 7.2

合并时间: 2026-07-23 08:30

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/31757>

执行摘要

- 一句话: Miles ROCm 构建迁移至 docker/build.py, 测试切换至 7.2
- 推荐动作: 值得精读, 尤其是关注 CI 安全实践 (最小权限原则) 和外部脚本集成的团队。设计权衡如测试覆盖与维护成本的取舍也值得参考。

功能与动机

根据 PR 描述, 原有内联 `docker build` 在各 workflow 中重复定义构建参数, 不易维护。通过调用 Miles 仓库统一的 `docker/build.py`, 所有 variant 构建参数集中存储在 Miles 的 `VARIANTS` 表中。同时将夜间测试迁移至 ROCm 7.2, 以跟上 AMD 最新驱动版本 (ROCm 7.0 镜像仍继续发布, 但不再有 nightly 测试覆盖)。

实现拆解

1. 修改 release workflow: 更新 `release-docker-amd-miles-rocm700-nightly.yml` 和 `release-docker-amd-miles-rocm720-nightly.yml`, 将矩阵从 `include` 格式简化为 `variant` 列表, 移除内联 `docker build`, 改为调用 `python3 docker/build.py --variant ... --image-tag custom --custom-tag miles`, 然后添加日期标签并推送。
2. 添加依赖和权限加固: 在 `build` 步骤前增加 `pip install typer`, 确保依赖可用; 添加 `persist-credentials: false` 和 `permissions: contents: read` 以限制凭证泄露风险。
3. 迁移夜间测试: 将 `nightly-test-amd-miles-rocm700.yml` 重命名为 `nightly-test-amd-miles-rocm720.yml`, 将所有 ROCm 7.0 引用改为 7.2, 更新镜像标签基准, 工作流名称改为“Nightly Test Miles (AMD ROCm 7.2)”。
4. 归档构建参数: 移除 `MILES_COMMIT` 构建参数, 仅在日志中记录 Miles commit SHA, 接受微小提交漂移。

关键文件:

- `.github/workflows/release-docker-amd-miles-rocm700-nightly.yml` (模块 CI/CD; 类别 infra; 类型 infrastructure): 核心变更文件, 展示从内联 `docker build` 到集中式 `build.py` 的迁移, 包含矩阵简化、依赖安装和权限加固。
- `.github/workflows/release-docker-amd-miles-rocm720-nightly.yml` (模块 CI/CD; 类别 infra; 类型 infrastructure): 与 `rocm700` 工作流镜像对称变更, 仅针对 ROCm 7.2 variant。

- `.github/workflows/nightly-test-amd-miles-rocm720.yml` (模块 CI/CD; 类别 infra; 类型 rename-or-move) : 重命名并修改为 ROCm 7.2 测试工作流, 决定覆盖迁移。

关键符号: 未识别

评论区精华

Review 中有几条重要讨论:

- P1 Typer 依赖: bingxche 指出缺少 typer 模块会导致 `ModuleNotFoundError`, 要求在 build 前安装。已通过增加 `pip install typer` 修复。
- P2 MILES_COMMIT 参数: bingxche 建议保留 SHA 以确保镜像内容可追溯, 但 XinyuJiangCMU 认为 Miles 更新频率低, 漂移可接受, 未采纳。
- P3 测试覆盖迁移: bingxche 询问放弃 ROCm 7.0 测试是否意图明确, XinyuJiangCMU 确认是故意之举, 并更新了 PR 描述。
- 凭证安全: bingxche 建议设置 `persist-credentials: false` 并添加 `permissions: contents: read` 以降低凭证风险, 已修复。
 - 安装 Typer 依赖 (correctness): XinyuJiangCMU 在第三个 commit 中添加了 `pip install typer`, 问题已解决。
 - MILES_COMMIT 构建参数移除 (design): 未采纳, 维持当前简化方案。
 - 测试迁移至 ROCm 7.2 (question): XinyuJiangCMU 确认是故意之举, 并更新了 PR 描述。
 - 凭证安全加固 (security): 已修复, 第三个 commit 添加了这些配置。

风险与影响

- 风险:
 - 外部脚本风险: Miles 的 `docker/build.py` 可能随机更新, 改变构建行为导致失败或错误配置。
 - 提交漂移: 不再传递 MILES_COMMIT, Dockerfile 可能克隆 Miles 的 main 分支而非 checkout 的 HEAD, 造成镜像内容与记录 SHA 不一致。
 - 测试覆盖丢失: ROCm 7.0 不再有 nightly 测试, 回归问题可能较晚发现。
 - 凭证仍暴露于外部代码: 虽然已加固, 但 Docker Hub 登录后 runner 仍会执行来自第三方仓库的 Python 脚本。
- 影响:
 - 用户: 无直接影响, 镜像标签格式保持不变。
 - 团队 CI 维护者: 需要了解 Miles 仓库的 `docker/build.py` 脚本参数和预期行为; 夜间测试仅覆盖 ROCm 7.2, 7.0 需手动触发。
 - 系统: 构建时间因依赖安装而微增, 但整体影响低。
 - 风险标记: 外部脚本依赖, 提交漂移未解决, ROCm 7.0 无 nightly 测试覆盖, 权限加固已修复

关联脉络

- 暂无明显关联 PR