

PR #7140 完整报告

verl-project/verl

[docker] chore: Rocm714 doc update

合并时间: 2026-07-24 23:40

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7140>

执行摘要

- 一句话: 更新 ROCm 7.14 Docker 构建文档与示例
- 推荐动作: 值得关注的设计决策是将主 Dockerfile 切换为基于预构建基镜像, 这显著降低了维护成本。建议精读 Dockerfile.rocm 和 new Dockerfile.rocm702 的结构差异以理解下一步构建策略。对于 AMD 用户, 应优先使用更新后的 amd_quick_start.rst 作为起点。

功能与动机

随着 ROCm 7.14 的发布, 需要更新 Docker 构建文档以使用最新的基镜像和软件版本, 并提供可复现的 Dockerfile 示例。同时清理已过时的旧文档, 降低用户混淆。

实现拆解

1. Dockerfile 架构调整: 将主 Dockerfile.rocm 重构为基于预构建的 rocm/primus:v26.4 基镜像, 避免从源头编译, 简化构建流程并提高可复现性。新增 Dockerfile.rocm702 作为历史版本 (ROCm 7.0.2) 的参考, 保留原有的从 ubuntu:22.04 开始的多阶段构建。
2. 知识库文档迁移: 删除旧的、内容庞大的 amd_build_dockerfile_page.rst (-796 行), 其内容已过时且覆盖了 ROCm 6.x 和手动编译细节; 将 amd_quick_start.rst 更新为当前推荐的快速入门指南, 更新软件版本号 (ROCm 7.14、PyTorch 2.12、vLLM 0.22.1rc1、SGLang 0.5.15 等), 并更新硬件支持列表 (MI350X 加入), 同时增加 SGLang 为完全支持。
3. 辅助文档同步: 更新 docker/rocm/README.md 以反映新的构建基镜像和版本清单; 修改 docs/amd_tutorial/index.rst 移除对已删除页面的引用, 并将入口统一指向 amd_quick_start.rst。

关键文件:

- docker/rocm/Dockerfile.rocm (模块 Docker 镜像; 类别 infra; 类型 infrastructure): 主 Docker 构建文件, 从 ROCm 7.0.2 的自定义编译模式切换为基于预构建基镜像 rocm/primus:v26.4, 显著简化构建流程。这是本次更新的核心变更。
- docker/rocm/Dockerfile.rocm702 (模块 Docker 镜像; 类别 infra; 类型 infrastructure): 新增的文件, 保留了原 ROCm 7.0.2 的多阶段构建 Dockerfile 作为历史参考, 使用户可以回溯旧版本构建方式。
- docs/amd_tutorial/amd_quick_start.rst (模块 AMD 教程; 类别 docs; 类型 documentation): 更新后的快速入门文档, 反映最新软件版本和硬件支持, 并新增

SGLang 支持声明。是用户查阅的主要入口。

- docs/amd_tutorial/amd_build_dockerfile_page.rst (模块 AMD 教程; 类别 docs; 类型 deletion) : 被删除的旧文档, 内容超过 700 行, 覆盖 ROCm 6.3 构建细节。删除表明文档体系向更简洁的 quick start 迁移。
- docker/rocm/README.md (模块 Docker 镜像; 类别 docs; 类型 documentation) : 同步更新版本信息和构建说明, 反映基镜像变更和组件版本升级。
- docs/amd_tutorial/index.rst (模块 AMD 教程; 类别 docs; 类型 documentation) : 移除对已删除文档的引用, 确保文档索引一致性。

关键符号: 未识别

评论区精华

在 Review 中, wuxibin89 询问是否可以删除 amd_build_dockerfile_page.rst。mingjielu 确认该文档是旧版 (ROCm 6.3) 构建指南, 现已由 amd_quick_start.rst 替代, 并已更新 README 和索引指向新文档。双方达成一致后文件被删除。结论: 删除旧文档, 保留历史 Dockerfile 作为参考。

- 删除旧文档 amd_build_dockerfile_page.rst (design): 一致同意删除, 旧文档已被新文档替代。

风险与影响

- 风险: 由于 Dockerfile.rocm 的构建方式从自适应编译改为基于预编译基镜像, 依赖链发生改变 (不再需要从头编译 Flash Attention、TE 等), 可能对希望自定义 GPU 架构或修改底层组件的用户造成影响。文档删除可能使依赖旧教程的用户暂时找不到信息, 但已有明确重定向。无代码执行风险。
- 影响: 核心影响是对 AMD GPU 用户 (MI300X/MI325X/MI350X/MI355X) : 构建时间大幅缩短, 但失去了一些可配置性; 文档结构更清晰, 入门门槛降低。对其他平台用户无影响。
- 风险标记: 构建模式切换, 文档重构

关联脉络

- PR #6388 Unknown (referenced in PR body): PR body 中提及作为相似 PR 的参考, 但具体信息未提供。