

PR #27555 完整报告

sgl-project/sglang

[AMD] update ROCm AITER commit

合并时间: 2026-06-08 23:49

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/27555>

执行摘要

- 一句话: 更新 ROCm Docker 镜像中 AITER 依赖版本
- 推荐动作: 此 PR 为常规的依赖版本更新, 无需深入代码审查。建议关注 AMD CI 结果, 确认 AITER 版本通过所有测试。如果后续出现与 AITER 相关的回归, 可以追溯到此 commit。

功能与动机

晋升 ROCm Docker 镜像中固定的 AITER commit, 以引入最新的 AITER 修复和性能改进。PR 描述中指出 'N/A - Docker build dependency bump only. Validated via AMD CI.', 表明此变更为纯粹的构建依赖升级, 无准确性或速度测试要求。

实现拆解

在 `docker/rocm.Dockerfile` 中, 将四个基础镜像阶段的 `AITER_COMMIT_DEFAULT` 环境变量值从旧 commit `6740fb7911db231e6eadad45194ad6713d25f9fb` 替换为新 commit `7a8ff7dd4ae3063ff1a18622a46460125c84370e`。每个镜像阶段对应不同的 ROCm 基础镜像配置 (如 `gfx942`、`gfx942-rocm720`、`gfx950`、`gfx950-rocm720`) , 但所有阶段均使用相同的 AITER commit。此变更通过 AMD CI 验证。

关键文件:

- `docker/rocm.Dockerfile` (模块 部署脚本; 类别 `infra`; 类型 `infrastructure`) : 唯一变更文件, 更新了所有 ROCm 基础镜像的 AITER commit hash, 直接影响 AMD ROCm Docker 镜像构建。

关键符号: 未识别

关键源码片段

`docker/rocm.Dockerfile`

唯一变更文件, 更新了所有 ROCm 基础镜像的 AITER commit hash, 直接影响 AMD ROCm Docker 镜像构建。

```
# 从旧 commit 6740fb79 更新为新 commit 7a8ff7dd
# 以下为四个不同 ROCm 基础镜像阶段的 AITER commit 变量设置
# Base image 942 without triton
```

```
ENV AITER_COMMIT_DEFAULT="7a8ff7dd4ae3063ff1a18622a46460125c84370e"  
# Base image 942 with rocm720 and triton  
ENV AITER_COMMIT_DEFAULT="7a8ff7dd4ae3063ff1a18622a46460125c84370e"  
# Base image 950 without triton  
ENV AITER_COMMIT_DEFAULT="7a8ff7dd4ae3063ff1a18622a46460125c84370e"  
# Base image 950 with rocm720 and triton  
ENV AITER_COMMIT_DEFAULT="7a8ff7dd4ae3063ff1a18622a46460125c84370e"
```

评论区精华

无人工 review 评论。只有 Gemini Code Assist bot 自动生成的代码审查，确认无问题。
HaiShaw 直接批准。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低。变更仅更新了构建时依赖的 commit 哈希，不涉及代码逻辑或运行时行为。但新 commit 可能引入新的 bug 或与现有 ROCm 环境的兼容性问题，不过已经通过 AMD CI 验证。如果新 AITER 版本存在回归，可能会导致 ROCm Docker 镜像构建失败或运行时性能下降。
- 影响：影响范围仅限于使用 ROCm Docker 镜像的用户和 AMD CI 流程。所有构建该镜像的环境（gfx942、gfx950 系列）都将使用新版本的 AITER。由于是依赖升级，对现有功能的直接影响为中等，但长期有助于获得修复和改进。
- 风险标记：依赖升级未经长期验证

关联脉络

- PR #27387 build(sgl-kernel): support configurable mirrors for restricted networks: 同为 sgl-kernel 构建基础设施变更，涉及 Dockerfile 和构建依赖管理。
- PR #26922 [PD][MoRI] Drive KV transfers with a sharded synchronous worker pool: 同为 AMD 相关的基础设施变更，修改了 docker/rocm.Dockerfile 和依赖管理。