

PR #32643 完整报告

sgl-project/sglang

[AMD] Add Kimi K3 ROCm 7.2 nightly image

合并时间: 2026-07-28 18:45

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/32643>

执行摘要

- 一句话: 新增 Kimi K3 ROCm 7.2 每日镜像发布任务
- 推荐动作: 值得关注但无需深入审查。该 PR 是纯 CI 配置变更, 逻辑清晰。建议后期确认 kimi-k3 分支合并后是否同步更新此 job 的 checkout 分支。

功能与动机

为支持 Kimi K3 模型在 MI35X 上的持续集成测试与用户试用, 需要从临时上游 **kimi-k3** 分支自动构建每日 ROCm 7.2 镜像, 同时保持现有 ROCm 每日构建流程不受影响。

实现拆解

1. 新增任务 **publish_k3_mi35x**: 在 .github/workflows/release-docker-amd-rocm720-nightly.yml 中添加新的 CI job, 条件为 schedule 事件或手动选择 publish-k3。
2. 分支与构建参数: 该任务明确 checkout kimi-k3 分支, 并通过 --build-arg SGL_BRANCH=kimi-k3 让 Dockerfile 从该分支拉取代码。
3. 镜像标签与推送: 构建后先推送到内部仓库 rocm/sgl-dev, 再 retag 为 lmsysorg/sglang-rocm:rocm720-mi35x-k3-YYYYMMDD 并推送到公开 Docker Hub。
4. 手动触发支持: 在 workflow_dispatch 的 options 中添加 publish-k3 选项, 方便人工触发独立 CI 运行。
5. 不修改现有流程: 原有的 publish、push_local_registry 等 job 保持原样, 确保现有每日构建不受影响。

关键文件:

- .github/workflows/release-docker-amd-rocm720-nightly.yml (模块 CI/CD; 类别 infra; 类型 infrastructure): 该文件是整个 PR 唯一变更的文件, 新增了独立的 Kimi K3 每日镜像构建与发布 job。

关键符号: 未识别

评论区精华

该 PR 无 review 评论, 仅由 bingxche 直接批准。PR body 中声明为 draft 且 blocked by #32621, 但最终仍被合并。

- 暂无高价值评论线程

风险与影响

- 风险：风险较低。新 job 独立于现有流程，不会影响现有每日镜像构建。但若 kimi-k3 分支代码不稳定或构建失败，可能导致该 job 失败，但不会阻塞其他 CI。此外，新增的 publish-k3 选项增加了 workflow_dispatch 的复杂度。
- 影响：影响范围限定在 CI/CD 流程。对用户而言，可获得每日更新的 Kimi K3 ROCm 7.2 镜像；对团队而言，维护人员需要关注该 job 的构建状态。影响程度较低，仅涉及基础设施。
- 风险标记：CI 配置变更，依赖临时分支

关联脉络

- PR #32621 [kimi-k3] (unrelated): PR body 声明本 PR blocked by #32621，需要该 PR 合入 kimi-k3 分支后此 CI 才能正常工作。
- PR #32542 docs(cookbook): add the Kimi-K3 serving cookbook: 该 PR 与 Kimi K3 相关，提供了部署配置等信息。