

PR #24155 完整报告

sgl-project/sglang

[AMD] Nightly image release for deepseek v4

合并时间: 2026-04-30 23:49

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/24155>

执行摘要

- 一句话: 添加 DeepSeek V4 AMD 夜间镜像发布
- 推荐动作: 这是一次纯粹的 CI/CD 配置变更, 对于非 AMD 平台的工程师可忽略。但 AMD 团队应关注构建成功率, 并考虑添加构建后的测试验证步骤。

功能与动机

PR body 提到目标为 'Release daily dsv4 image on branch amd/deepseek_v4', 为 DeepSeek V4 模型提供 AMD 平台每日构建的 Docker 镜像, 支持持续集成和测试。

实现拆解

1. 在工作流文件 `release-docker-amd-rocm720-nightly.yml` 中添加 `workflow_dispatch` 的输入 `job_select`, 支持选择 'publish_dsv4' 单独触发。
2. 在现有 `publish` 作业的条件中添加 `job_select` 检查, 允许选择性运行。
3. 新增 `publish_dsv4` 作业, 从 `amd/deepseek_v4` 分支检出代码, 获取版本号, 构建并推送镜像到 `rocm/sgl` 仓库, 镜像标签格式为 `rocm720-{gpu_arch}-{date}-{commit}-DSv4`。
4. 保留现有 `publish` 作业不变, 确保向后兼容。

关键文件:

- `.github/workflows/release-docker-amd-rocm720-nightly.yml` (模块 部署脚本; 类别 `infra`; 类型 `infrastructure`): 仅修改的文件, 添加了 `publish_dsv4` 作业和 `job_select` 输入支持

关键符号: 未识别

评论区精华

PR 未收到实质性审查讨论, 仅由合并者自行 approve。bot 评论涉及 `human-eval` 子模块, 但与本次变更无关。

- 暂无高价值评论线程

风险与影响

- 风险：主要风险是镜像构建失败或不稳定，可能阻塞 deepseek v4 的测试流程。但 workflow 仅在手动触发或定时触发时运行，不影响核心服务。需要确保 amd/deepseek_v4 分支的稳定性。
- 影响：对用户：AMD DeepSeek V4 用户将获得每日构建的 Docker 镜像。对系统：增加一个 CI/CD 作业，但运行频率低（每天一次或手动）。对团队：AMD 平台维护者需要维护镜像构建脚本。
- 风险标记：构建失败风险，分支依赖

关联脉络

- 暂无明显关联 PR