

PR #23778 完整报告

sgl-project/sglang

ci(deepseek-v4): add b300/grace-blackwell dev-branch build options

合并时间: 2026-04-27 07:44

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/23778>

执行摘要

- 一句话: 新增 DeepSeek-V4 开发分支 Docker 构建选项
- 推荐动作: 值得合并, 变更清晰且经过充分测试计划。建议其他项目组参考此模式 (通过 branch 字段和 tag 后缀实现多分支镜像构建)。

功能与动机

团队需要从 `deepseek_v4_dev` 分支发布 Docker 镜像, 用于 B300 和 Grace Blackwell 平台的开发 / 测试, 同时保持 `deepseek_v4` 稳定分支的构建不受干扰。PR body 明确指出: “This lets us release images from the `deepseek_v4_dev` branch without disturbing the stable `deepseek_v4` images.”

实现拆解

1. 新增 `workflow_dispatch` 输入: 在 `.github/workflows/release-docker-deepseek-v4.yml` 的 `on.workflow_dispatch.inputs` 下添加 `build_b300_dev` 和 `build_grace_blackwell_dev`, 类型均为 `boolean`, 默认 `true`。
2. 传递输入到矩阵生成脚本: 将新输入通过 `env` 映射为 `BUILD_B300_DEV` 和 `BUILD_GRACE_BLACKWELL_DEV` 环境变量, 供后续 `bash` 脚本使用。
3. 扩展矩阵条目生成逻辑: 在 `jobs.setup.outputs.matrix` 的 `bash` 脚本中增加两个 `if` 分支, 当 `BUILD_B300_DEV=true` 或 `BUILD_GRACE_BLACKWELL_DEV=true` 时, 向 `entries` 数组添加对应 JSON 对象。每个对象新增 `branch` 字段: 稳定分支使用 `deepseek_v4`, 开发分支使用 `deepseek_v4_dev`。
4. 修改 `checkout` 步骤: 将 `actions/checkout@v4` 的 `ref` 从硬编码的 `deepseek_v4` 改为 `${{matrix.branch}}`, 使每个矩阵任务能检出对应分支的代码。
5. 镜像标签区分: 开发分支的镜像 tag 后缀为 `-dev` (如 `deepseek-v4-b300-dev`), 避免与稳定镜像冲突。

关键文件:

- `.github/workflows/release-docker-deepseek-v4.yml` (模块部署脚本; 类别 `infra`; 类型 `infrastructure`): 唯一变更文件, 新增 `workflow_dispatch` 输入、矩阵生成逻辑和动态 `checkout` 分支支持。

关键符号: 未识别

评论区精华

该 PR 无 review 评论，仅有一条 Gemini 自动回复表示无法生成 review。讨论亮点主要来自 PR body 中的测试计划，预期验证：

- 只触发 `build_b300_dev=true` 时，矩阵应只生成一个 `deepseek-v4-b300-dev` 条目并检出 `deepseek_v4_dev`。
- 只触发 `build_grace_blackwell_dev=true` 时，ARM runner 应检出 `deepseek_v4_dev` 并推送 `deepseek-v4-grace-blackwell-dev`。
- 触发原有构建（如 `build_b300=true`）时，行为不变。
- 暂无高价值评论线程

风险与影响

- 风险：风险较低：
 - 新输入默认值为 `true`，若用户不熟悉可能意外触发开发分支构建，但不会影响稳定镜像（tag 不同）。
 - `checkout` 步骤引用 `matrix.branch`，若 `branch` 字段缺失（如旧版 workflow 未升级），`actions/checkout` 可能报错。但本 PR 同步修改了所有矩阵条目以保证字段存在。
 - 实际构建成功与否取决于 `deepseek_v4_dev` 分支上 `Dockerfile` 的可用性。
- 影响：影响范围限于 CI/CD 流程：
 - 为 DeepSeek-V4 开发团队提供独立的 Docker 构建通道，加速开发和测试迭代。
 - 稳定分支构建不受影响，tag 隔离确保镜像拉取无冲突。
 - workflow 复杂度略有增加（多两个输入参数），但维护成本低。
 - 风险标记：缺少测试覆盖

关联脉络

- PR #23728 `ci: add docker release workflow for deepseek_v4 branch`: 本 PR 在此 workflow 基础上扩展，新增 `dev` 分支构建选项。
- PR #23730 `[CI] release-docker-deepseek-v4: select which flavors to push`: 同一 workflow 的早期优化，引入分支选择逻辑，本 PR 进一步泛化。