

PR #32596 完整报告

sgl-project/sglang

sgl-kernel-npu tag update to 2026.7.27

合并时间: 2026-07-29 10:22

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/32596>

执行摘要

- 一句话: 更新 sgl-kernel-npu 标记版本
- 推荐动作: 可通过, 属于常规依赖版本升级。建议在合并前确认新包在目标硬件上通过 CI 测试。

功能与动机

sgl-kernel-npu 的 .whl 包已打补丁以适配不同机器的 A2 包, 因此需要将标签更新至 2026.7.27。

实现拆解

1. 在 .github/workflows/release-docker-npu.yml 中将 SGLANG_KERNEL_NPU_TAG 构建参数从 2026.7.2 改为 2026.7.27。
2. 在 .github/workflows/release-docker-npu-nightly.yml 中做相同更改。
3. 在 scripts/ci/npu/npu_ci_install_dependency.sh 中更新安装脚本的 SGLANG_KERNEL_NPU_TAG 环境变量, 以从正确的 GitHub Release 下载 wheel 包。

关键文件:

- .github/workflows/release-docker-npu-nightly.yml (模块 CI; 类别 infra; 类型 infrastructure) : 更新夜间 Docker 构建中的 NPU 内核标签。
- .github/workflows/release-docker-npu.yml (模块 CI; 类别 infra; 类型 infrastructure) : 更新稳定 Docker 构建中的 NPU 内核标签。
- scripts/ci/npu/npu_ci_install_dependency.sh (模块 CI; 类别 infra; 类型 infrastructure) : 更新 CI 安装脚本中的 NPU 内核版本变量。

关键符号: 未识别

关键源码片段

[.github/workflows/release-docker-npu.yml](#)

更新稳定 Docker 构建中的 NPU 内核标签。

片段: 构建参数块

build-args: |

SGLANG_KERNEL_NPU_TAG=2026.7.27 # 更新标签以使用打补丁后的 wheel 包

```
CANN_VERSION=${{ matrix.cann_version }}  
DEVICE_TYPE=${{ matrix.device_type }}  
SGLANG_TAG=${{ steps.version.outputs.version }}
```

评论区精华

无讨论。

- 暂无高价值评论线程

风险与影响

- 风险：低风险：仅替换版本标签字符串，不涉及逻辑变更。但若新发布包存在兼容性问题，则需回滚。
- 影响：影响限于 NPU 相关构建和 CI 流程，使 Docker 镜像及 CI 环境使用最新的 sgl-kernel-npu 版本。
- 风险标记：依赖升级

关联脉络

- 暂无明显关联 PR