

# PR #25730 完整报告

sgl-project/sglang

[XPU] upgrade triton-xpu version to 3.7.1

合并时间: 2026-05-21 10:29

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/25730>

## 执行摘要

- 一句话: 升级 triton-xpu 至 3.7.1
- 推荐动作: 该 PR 是简单的依赖升级, 无需精读。可关注的点: 正式索引和测试索引的 URL 变化, 以及版本锁定的实践, 适用于其他依赖管理场景。

## 功能与动机

triton-xpu 3.7.1 是正式发布版本, 而此前使用的版本来自测试索引且未锁定版本号。升级到正式版可避免 torch 2.11.0+xpu 打包版本中的已知 bug, 并提供更稳定的执行环境。

## 实现拆解

1. 更新 Docker 构建文件(docker/xpu.Dockerfile): 将 `pip install triton-xpu --index-url https://download.pytorch.org/whl/test/xpu --force-reinstall` 替换为 `pip install triton-xpu==3.7.1 --index-url https://download.pytorch.org/whl/xpu --force-reinstall`, 即锁定版本并切换到正式 PyTorch XPU 仓库。
2. 更新安装文档(docs\_new/docs/hardware-platforms/xpu.mdx): 同样将 pip 安装命令中的 triton-xpu 版本锁定为 3.7.1, 并将索引 URL 从测试通道改为正式通道, 同时移除注释 `# index will change after torch 2.12 release`。

关键文件:

- docker/xpu.Dockerfile (模块 部署脚本; 类别 infra; 类型 infrastructure): 修改了 triton-xpu 的安装指令, 锁定版本并切换索引, 直接影响 XPU Docker 镜像的构建过程。
- docs\_new/docs/hardware-platforms/xpu.mdx (模块 文档; 类别 other; 类型 core-logic): 更新了安装文档中的 triton-xpu 安装命令, 与 Dockerfile 保持一致, 引导用户使用正式版本。

关键符号: 未识别

## 评论区精华

PR 无 review 讨论; 作者 @jiayisunx 请求 @Xia-Weiwen 和 @mingfeima 审核, 两人均未发表评论, @Xia-Weiwen 直接批准。所有对话均为 GitHub bot 的配额警告或作者的审核请求, 无技术讨论。

- 暂无高价值评论线程

## 风险与影响

- 风险：风险极低。版本升级是向后兼容的，且仅有 2 个文件的小改动，测试索引与正式索引的差异很小。但由于 triton-xpu 3.7.1 可能与某些旧版 PyTorch 或特定 Intel GPU 驱动不兼容，需要确保 XPU 构建环境满足新版本要求。
- 影响：影响范围限于 Intel GPU (XPU) 平台的用户。从测试索引切换到正式索引意味着用户将获得更稳定的 triton-xpu 版本，但可能不再包含某些测试中的新特性或修补。由于版本已锁定，构建结果可重复。
- 风险标记：依赖版本升级

## 关联脉络

- 暂无明显关联 PR