

PR #26053 完整报告

sgl-project/sglang

[XPU] add apache-tvm-ffi dependency

合并时间: 2026-05-22 16:09

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/26053>

执行摘要

- 一句话: XPU 安装文档与 Dockerfile 新增 tvn-ffi 依赖
- 推荐动作: 该 PR 为常规依赖补充, 无需精读。但可关注后续 XPU 构建是否因版本冲突出现问题。

功能与动机

XPU 环境缺少 tvn_ffi 模块, 导致运行时异常; PR body 明确说明 'Install apache-tvm-ffi on XPU (Intel GPU) to resolve missing tvn_ffi module.'

实现拆解

1. 修改 XPU Dockerfile (`docker/xpu.Dockerfile`): 在 pip install 链中新增一行 pip install apache-tvm-ffi &&, 确保容器构建时自动安装该依赖。
2. 更新 XPU 安装文档 (`docs_new/docs/hardware-platforms/xpu.mdx`): 在 pip 安装步骤后添加注释 # xgrammar requires apache-tvm-ffi 并补全安装命令, 指导用户手动安装。
3. 无测试、配置或部署配套变更。

关键文件:

- `docker/xpu.Dockerfile` (模块 部署脚本; 类别 infra; 类型 infrastructure): XPU Docker 构建的入口文件, 新增 tvn-ffi 安装命令确保容器环境下依赖可用。
- `docs_new/docs/hardware-platforms/xpu.mdx` (模块 文档; 类别 other; 类型 core-logic): XPU 平台安装文档, 告知用户手动安装 tvn-ffi。

关键符号: 未识别

关键源码片段

`docker/xpu.Dockerfile`

XPU Docker 构建的入口文件, 新增 tvn-ffi 安装命令确保容器环境下依赖可用。

```
# docker/xpu.Dockerfile (片段)
RUN --mount=type=secret,id=github_token \
  cp pyproject_xpu.toml pyproject.toml && \
  pip install . --extra-index-url https://download.pytorch.org/whl/xpu && \
  pip install --no-deps xgrammar==0.1.33 && \
```

```
pip install apache-tvm-ffi && \ # 新增: 安装 apache-tvm-ffi 以提供 tvm_ffi 模块
pip install triton-xpu==3.7.1 --index-url https://download.pytorch.org/whl/xpu --force-reinstall
&& \
echo ". /home/sdp/miniforge3/bin/activate; conda activate py${PYTHON_VERSION}; cd /home/
sdp" >> /home/sdp/.bashrc
```

评论区精华

无 review 讨论评论。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：仅新增依赖安装步骤，不修改任何运行时逻辑。潜在风险包括 tvm-ffi 版本兼容性问题或安装失败导致流水线中断，但可通过锁定版本缓解。
- 影响：影响范围仅限于 XPU（Intel GPU）平台的用户和 CI 构建。对现有 CUDA/AMD 等其他平台无影响。用户需确保网络可访问 PyPI 上的 apache-tvm-ffi 包。
- 风险标记：缺少版本锁定，无测试覆盖

关联脉络

- 暂无明显关联 PR