

PR #29676 完整报告

sgl-project/sglang

[Docs] Add --prerelease=allow so uv installs the latest sglang

合并时间: 2026-06-30 05:13

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/29676>

PR 分析报告

执行摘要

本次 PR 在 sglang 安装文档中对 `uv pip install sglang` 命令添加了 `--prerelease=allow` 参数，解决 uv 因默认拒绝预发布依赖（`flash-attn-4`）而回溯安装旧版本 sglang 的问题。变更仅涉及两处文档文件，风险极低，但能确保 uv 用户安装最新版本，与 pip 行为对齐。

功能与动机

用户按照文档推荐的 `uv pip install sglang` 命令会静默安装 sglang 0.5.9 而非最新版本。根本原因是 sglang 0.5.10+ 依赖 `flash-attn-4` (`>=4.0.0b4 / >=4.0.0b9`)，而 `flash-attn-4` 在 PyPI 上仅发布预发布版本 (`4.0.0b3 ... 4.0.0b19`)，uv 默认拒绝预发布版本，导致回溯至最后一个不需要 `flash-attn-4` 的 0.5.9 版本。而 pip 接受预发布版本所以正常安装最新版。

实现拆解

1. 在 docs_new/docs/get-started/install.mdx 中修改两处命令：默认安装块和 CUDA 12 安装块。
2. 在 docs_new/docs/get-started/quickstart.mdx 中修改一处命令：Installation (Pip / uv) Tab 下的 uv 命令。
3. 验证结果：添加 `--prerelease=allow` 后，`uv pip install sglang` 解析为 `sglang==0.5.13.post1`（附带 `flash-attn-4==4.0.0b19` 等依赖）。

无源码变更，仅文档 Markdown 文件修改。

评论区精华

无 review 讨论。直接审核批准。

风险与影响

- 风险：极低。仅修改文档命令参数，不影响任何代码逻辑。
- 影响：使用 uv 安装 sglang 的用户将能通过文档推荐的命令获得最新版本，避免版本回溯；与 pip 用户的体验对齐。

关联脉络

未发现直接关联的 Issue 或其他 PR。属于独立文档修复。