

PR #35615 完整报告

sgl-project/sglang

[diffusion] ci: use canonical residency selector

合并时间: 2026-08-20 10:50

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/35615>

执行摘要

- 一句话: FLUX 对比 CI 改用 canonical residency 选择器
- 推荐动作: 值得快速浏览, 作为 diffusion nightly CI 配置维护的参考。可关注 canonical --component-residency 参数在 CI 中的实际验证情况, 以确认其与 legacy 参数等价。

功能与动机

PR 描述指出, FLUX 对比任务仍通过兼容性 flag `--dit-layerwise-offload false` 表达 resident DiT 放置, 而 nightly 配置应使用 canonical residency selector (`--component-residency dit=resident`), 同时 legacy CLI 路径需继续支持用户。

实现拆解

1. 修改对比配置: 在 `scripts/ci/utils/diffusion/comparison_configs.json` 中, 将 FLUX.1 与 FLUX.2 的 `serve_args` 从 `--dit-layerwise-offload false` 改为 `--component-residency dit=resident`, 其余参数不变。
2. 保持向后兼容: 未删除或修改 any legacy CLI 处理逻辑, 确保旧参数仍可用。
3. 验证方式: 依赖现有 nightly 任务覆盖此配置, 未本地运行准确性测试, 并请求 NVIDIA CI 验证 canonical argument path。

关键文件:

- `scripts/ci/utils/diffusion/comparison_configs.json` (模块 CI 配置; 类别 infra; 类型 infrastructure): 核心配置文件, 替换 FLUX.1 与 FLUX.2 的 residency 参数, 使 CI 配置与文档一致。

关键符号: 未识别

评论区精华

无 review 讨论。

- 暂无高价值评论线程

风险与影响

- 风险：由于未进行本地准确性测试，存在新参数路径解析或行为与预期不符的潜在风险。但改动仅涉及 CI 配置，不触及源码或运行时逻辑，且体现实测将验证新旧参数等价，风险较低。
- 影响：影响范围限于 CI nightly 对比测试的配置，预期行为无变化。改善配置可维护性，并验证 canonical residency selector 在实际任务中的可用性，为未来迁移提供依据。团队需留意 CI 结果以确保新路径正常工作。
- 风险标记：CI 配置变更，未运行准确性测试

关联脉络

- PR #35538 [diffusion] fix: stop reserving NCCL device buffers for single-rank groups: 同为 diffusion 模块的 CI/ 运行时配置调整，涉及 diffusion nightly 测试相关基础设施。
- PR #35587 docker: fix CUDA-13 build — rename NCCL_VERSION ARG to avoid base image ENV collision: 同为基础设施修复，关联 diff 模块的 CI 稳定性。