

PR #24400 完整报告

sgl-project/sclang

[diffusion] CI: pin diffusion consistency GT revision

合并时间: 2026-05-05 11:53

原文链接: <http://prhub.com.cn/sgl-project/sclang/pull/24400>

执行摘要

- 一句话: 固化 diffusion 一致性测试 GT 来源版本
- 推荐动作: 值得精读, 作为测试基础设施最佳实践示例——通过 pinned revision 消除外部依赖波动。

功能与动机

之前 diffusion 一致性检查使用的 GT 数据来自 `ci-data/main` 分支, `ci-data` 的非关联更新可能导致 PR 测试行为变化。PR body 明确说明: "Binding the default GT source to a concrete ci-data commit makes the check reproducible."

实现拆解

1. 定义 pinned revision 常量: 在 `test_utils.py` 中新增 `SGL_TEST_FILES_CI_DATA_REVISION` 常量, 值为 `437539b330592b1421d239b22d7d76b4a6d08fda`。
2. 重构 GT URL 构造: 将原来的直接硬编码 URL 替换为基于 `SGL_TEST_FILES_CONSISTENCY_GT_ROOT` 动态拼接, 根路径中嵌入 revision 值, 确保所有 GT URL 指向固定 commit。
3. 新增断言测试: 在 `test_consistency_metrics.py` 中新增 `test_consistency_gt_urls_are_pinned_to_ci_data_revision`, 断言 `SGL_TEST_FILES_CONSISTENCY_GT_ROOT` 不含 `/ci-data/main/`, 且两个 GT base URL 均包含 revision 路径。
4. 增强缺失 GT 错误信息: 在 `test_server_common.py` 的 GT 丢失提示中增加 Pinned revision used by this check: {`SGL_TEST_FILES_CI_DATA_REVISION`}, 帮助开发者定位版本。

关键文件:

- `python/sclang/multimodal_gen/test/test_utils.py` (模块 测试工具; 类别 test; 类型 test-coverage; 符号 `SGL_TEST_FILES_CI_DATA_REVISION`, `SGL_TEST_FILES_CONSISTENCY_GT_ROOT`, `SGL_TEST_FILES_OFFICIAL_CONSISTENCY_GT_BASE`, `SGL_TEST_FILES_SGLANG_CONSISTENCY_GT_BASE`): 定义 pinned revision 常量并重构 GT URL 构造方式, 是本次变更的核心配置所在。

- python/sglang/multimodal_gen/test/test_consistency_metrics.py (模块 一致性测试; 类别 test; 类型 test-coverage; 符号 test_consistency_gt_urls_are_pinned_to_ci_data_revision) : 新增断言测试, 确保 GT URL 不再指向 main, 是变更的验证环节。
- python/sglang/multimodal_gen/test/server/test_server_common.py (模块 服务测试; 类别 test; 类型 test-coverage) : 在缺失 GT 的错误信息中暴露 pinned revision, 提升可调试性。

关键符号: test_consistency_gt_urls_are_pinned_to_ci_data_revision

关键源码片段

python/sglang/multimodal_gen/test/test_utils.py

定义 pinned revision 常量并重构 GT URL 构造方式, 是本次变更的核心配置所在。

```
# 新增 pinned revision 常量
SGL_TEST_FILES_CI_DATA_REVISION = (
    "437539b330592b1421d239b22d7d76b4a6d08fda"
)

# 用 f-string 构造固定 revision 的 GT 根 URL, 不再使用 /ci-data/main/
SGL_TEST_FILES_CONSISTENCY_GT_ROOT = (
    "https://raw.githubusercontent.com/"
    f"sgl-project/ci-data/{SGL_TEST_FILES_CI_DATA_REVISION}/"
    "diffusion-ci/consistency_gt"
)

# 两个 base URL 从根 URL 派生, 确保所有 GT 数据来源固定
SGL_TEST_FILES_OFFICIAL_CONSISTENCY_GT_BASE = (
    f"{SGL_TEST_FILES_CONSISTENCY_GT_ROOT}/official_generated"
)
SGL_TEST_FILES_SGLANG_CONSISTENCY_GT_BASE = (
    f"{SGL_TEST_FILES_CONSISTENCY_GT_ROOT}/sglang_generated"
)
```

python/sglang/multimodal_gen/test/test_consistency_metrics.py

新增断言测试, 确保 GT URL 不再指向 main, 是变更的验证环节。

```
def test_consistency_gt_urls_are_pinned_to_ci_data_revision():
    # 构造期望的 revision 路径片段
    revision_path = (
        f"/ci-data/{test_utils.SGL_TEST_FILES_CI_DATA_REVISION}/"
    )
    # 确保根 URL 不再指向浮动的 main
    assert (
        "/ci-data/main/" not in test_utils.SGL_TEST_FILES_CONSISTENCY_GT_ROOT
    )
    # 确保两个 GT base URL 都包含了 pinned revision
    assert revision_path in test_utils.SGL_TEST_FILES_OFFICIAL_CONSISTENCY_GT_BASE
    assert revision_path in test_utils.SGL_TEST_FILES_SGLANG_CONSISTENCY_GT_BASE
```

python/sglang/multimodal_gen/test/server/test_server_common.py

在缺失 GT 的错误信息中暴露 pinned revision，提升可调试性。

```
# 在导入处新增 SGL_TEST_FILES_CI_DATA_REVISION
from sglang.multimodal_gen.test.test_utils import (
    SGL_TEST_FILES_CI_DATA_REVISION,
    # ... 其他原有导入
)

# 在 GT 缺失的 logger.error 消息中加入 revision 信息
logger.error(f"""
...
Repository: https://github.com/sgl-project/ci-data (path: diffusion-ci/consistency_gt/sglang_
generated/)
Pinned revision used by this check: {SGL_TEST_FILES_CI_DATA_REVISION}
...
""")
```

评论区精华

本 PR 无 review 评论，变更直接合并。唯一机器人评论提到每日配额限制，未涉及技术讨论。

- 暂无高价值评论线程

风险与影响

- 风险：该 PR 仅修改测试基础设施常量和新增测试，不涉及任何核心逻辑或运行时路径。风险极低，但需确保 ci-data 仓库 437539b330592b1421d239b22d7d76b4a6d08fda 提交长期存在，否则 GT URL 会永久失效。
- 影响：影响范围仅限 diffusion 模块 CI 测试。变更后，一致性检查结果不再受 ci-data 仓库 main 分支变动影响，测试可复现性增强。新增的断言测试会在 CI 中阻止未来引入指向 main 的 URL。
- 风险标记：外部依赖固定值可能过期

关联脉络

- 暂无明显关联 PR