

PR #21271 完整报告

sgl-project/sclang

Fix disaggregation test bootstrap port conflict

合并时间: 2026-03-25 12:14

原文链接: <http://prhub.com.cn/sgl-project/sclang/pull/21271>

执行摘要

- 一句话: 修复 disaggregation 测试中的 bootstrap 端口冲突, 避免 CI 失败。
- 推荐动作: 对于测试维护者和 CI 工程师, 建议快速浏览以了解端口配置模式; 对于其他开发者, 变更简单, 无需深度关注。

功能与动机

PR 标题和 body 中的 GitHub Actions 链接 (<https://github.com/sgl-project/sclang/actions/runs/23436739759/job/68185165551#step:6:4403>) 表明, CI 测试失败是由于 bootstrap 端口冲突导致的。作者 ispobock 通过此修复来避免测试中断, 提升开发流程效率。

实现拆解

实现方案分为两个层次: 首先, 在测试夹具文件 `python/sclang/test/server_fixtures/disaggregation_fixture.py` 中, 修改 `setUpClass` 方法, 添加 `cls.bootstrap_port = f"{int(base_port) + 500}"`, 并更新打印语句以包含该端口。其次, 在 10 个 disaggregation 测试文件 (如 `test_disaggregation_basic.py` 等) 中, 为 `start_prefill` 和 `start_decode` 方法添加命令行参数 `--disaggregation-bootstrap-port` 并引用 `cls.bootstrap_port`, 确保所有测试进程使用一致的端口配置。

关键文件:

- `python/sclang/test/server_fixtures/disaggregation_fixture.py` (模块 `test fixtures`): 定义了 `bootstrap` 端口变量, 是测试夹具的核心更改, 确保所有测试使用统一端口。
- `test/registered/disaggregation/test_disaggregation_basic.py` (模块 `disaggregation tests`): 作为示例, 展示了如何在 `prefill` 和 `decode` 启动命令中添加 `--disaggregation-bootstrap-port` 参数, 覆盖多个测试场景。

关键符号: `setUpClass`, `start_prefill`, `start_decode`

评论区精华

Review 过程简单, 仅有 hnyls2002 的批准, 无实质性讨论或争议。这表明变更被认可为直白的 bugfix。

- Approval (other): 变更被接受并合并。

风险与影响

- 风险：风险较低：变更仅限于测试代码，不影响生产环境。潜在风险包括端口计算可能与其他测试冲突（但基于现有端口偏移逻辑，风险可控），以及可能遗漏某些测试文件（但 PR 覆盖了所有相关的 disaggregation 测试文件）。
- 影响：影响范围有限：主要提升 CI 测试的稳定性和可靠性，减少因端口冲突导致的 flakiness，从而加速开发迭代。对最终用户或系统性能无直接影响。
- 风险标记：端口配置依赖，测试覆盖调整

关联脉络

- PR #21305 Increase flush cache timeout in hicache CI: 同样涉及 CI 测试稳定性优化，通过调整超时机制修复 flakiness。
- PR #21349 [CI] Reduce session correctness test to 30 turns to fix flakiness: 另一个测试修复 PR，通过减少测试轮数解决不稳定问题，与本 PR 同属提升 CI 可靠性的努力。