PR #25203 完整报告

sgl-project/sglang

ci: B200 conditional split + LPT_SLOP removal (stage-c partition 8→3)

合并时间：2026-05-14 09:40

原文链接：http://prhub.com.cn/sgl-project/sglang/pull/25203

执行摘要

- 一句话：B200 测试移至 nightly 并精简 CI 分区配置
- 推荐动作：推荐关注 CI 效率优化的团队阅读此 PR，特别是 compute_partitions.py 中的改动，展示了如何通过数据驱动校准和移除过度保护来精简 CI 配置。未来可考虑将更多非门控测试移至条件触发或 nightly 队列。

功能与动机

B200 测试在 per-commit 中运行了 8 个分区，但它们并不门控大多数典型 PR，且 est_time 被过度配置了 30-60%。通过移除非必要测试和移除冗余的 LPT_SLOP 填充，可以减少 CI 资源消耗，缩短开发者等待时间。

实现拆解

1. 测试套件迁移：在 7 个 B200 测试文件中，将 register_cuda_ci 的参数从 stage="stage-c", runner_config="4-gpu-b200" 替换为 suite="nightly-4-gpu-b200", nightly=True，并调低 est_time 至实际观测值（如 LoRA 测试从 300 秒降至 90-100 秒）。
2. LPT_SLOP 移除：在 scripts/ci/utils/compute_partitions.py 中删除 LPT_SLOP = 1.15 及其使用，分区数直接通过 max(1, math.ceil(total / TARGET_SECONDS)) 计算，转而依赖作业级 30 分钟超时和 MAX_PARTITION_SECONDS 检查作为安全网。
3. est_time 重新校准：更新剩余 B200 测试的 est_time，例如 test_gpt_oss_4gpu.py 的 B200 项从 740 秒降至 350 秒，test_nvidia_nemotron_3_super_nvfp4.py 从 710 秒降至 540 秒，使其更贴近实际运行时间。

关键文件：

- scripts/ci/utils/compute_partitions.py（模块 CI 脚本；类别 infra；类型 infrastructure；符号 compute_partitions）：移除 LPT_SLOP 常量并简化分区计算逻辑，是 CI 分区优化的核心变更。
- test/registered/lora/test_lora_gpt_oss_20b_logprob_diff.py（模块 LoRA 回归；类别 test；类型 test-coverage）：代表 7 个从 per-commit 移至 nightly 的 B200 LoRA 测试之一，展示了注册参数的变化。
- test/registered/4-gpu-models/test_gpt_oss_4gpu.py（模块 模型测试；类别 test；类型 test-coverage）：B200 测试的 est_time 从 740 秒降至 350 秒，基于实际观测进行校准。

关键符号：compute_partitions

关键源码片段

scripts/ci/utils/compute_partitions.py

移除 LPT_SLOP 常量并简化分区计算逻辑，是 CI 分区优化的核心变更。

```
defcompute_partitions(tests, full_parallel=False):     # ...     if suite in _STAGE_A_OVERRIDES:         size = _STAGE_A_OVERRIDES[suite]         max_parallel = size     else:         # 之前使用 LPT_SLOP = 1.15 进行填充，现直接除以 TARGET_SECONDS（20 分钟）         # 作业级 30 分钟超时作为最终安全网         size = max(1, math.ceil(total / TARGET_SECONDS))         max_parallel = size if full_parallel else compute_max_parallel(size)     # 检查平均时间是否超过硬性上限     if total / size > MAX_PARTITION_SECONDS:         raise RuntimeError(             f"Suite{suite!r}: total est_time{total:.0f}s / size{size}"             f"={total/size:.0f}s >{MAX_PARTITION_SECONDS}s"         ) _ 注：TARGET_SECONDS=1200，MAX_PARTITION_SECONDS=1800。_
```

test/registered/lora/test_lora_gpt_oss_20b_logprob_diff.py

代表 7 个从 per-commit 移至 nightly 的 B200 LoRA 测试之一，展示了注册参数的变化。

```
# 之前：register_cuda_ci(est_time=300, stage="stage-c", runner_config="4-gpu-b200")
# 之后：移至 nightly 套件，est_time 降至 90 秒
register_cuda_ci(
    est_time=90,
    suite="nightly-4-gpu-b200",
    nightly=True,
)
```

评论区精华

该 PR 未收到审核评论，仅包含作者触发的 rerun 请求和自动额度提示。由于变更属于纯基础设施优化且逻辑清晰，可能已通过内部对齐后直接合并。

- 暂无高价值评论线程

风险与影响

- 风险：主要风险有两方面：一是将测试移至 nightly 后会延迟 B200 相关问题的发现，但这些测试原本就不常触发失败，且 nightly 运行仍能覆盖；二是移除 LPT_SLOP 后，在极端不平衡场景下可能导致单个分区略超预期，但作业级 30 分钟硬超时和 MAX_PARTITION_SECONDS 检查提供了保护，风险较低。另外，est_time 重新校准基于单次观测，若模型更新或环境变化导致运行时间增加，可能需要再次校准。
- 影响：对开发者而言，CI 反馈速度提升：stage-c-test-4-gpu-b200 从 8 分区（~16 分钟）缩减至 3 分区（~19 分钟），相同时间窗内可运行更多 PR；其他套件如 stage-b-test-1-gpu-large 从 14 分区降至 12 分区，stage-c-test-8-gpu-h200 从 6 分区降至 5 分区。负面影响是 B200 相关的回归测试反馈延迟至 nightly 运行。

- 风险标记：CI 配置逻辑变更，测试覆盖降级，跨套件影响

关联脉络

- PR #25193 ci: compute matrix partition counts from est_time: 同样涉及 compute_partitions.py 的分区动态计算，本 PR 在其基础上移除了 LPT_SLOP。
- PR #25197 ci: decouple stage and runner for cuda registry: 涉及 CI 注册体系解耦，为本 PR 引入 suite 和 nightly 参数提供了基础设施。