

PR #22461 完整报告

sgl-project/sglang

[CI] Add GB200 nightly perf regression pipeline

合并时间: 2026-04-11 06:12

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/22461>

执行摘要

本 PR 新增了一个针对 GB200 集群的夜间性能回归测试管道，通过 GitHub Actions 和 Slurm 集成自动化运行 DeepSeek-R1 推理基准。该管道旨在监控高性能硬件上的性能稳定性，减少手动测试负担，并已通过 review 讨论优化了安全性和可靠性。

功能与动机

为解决性能回归监控的需求，PR body 明确指出“添加一个夜间 CI 管道，在 GB200 集群上运行 DeepSeek-R1 推理基准测试”。动机是确保系统在 NVIDIA GB200 等高端硬件上的表现一致，避免因代码变更引入性能下降。

实现拆解

实现分为多个模块：

- workflow 定义: `.github/workflows/nightly-perf-regression-gb200.yml` 文件定义了每日 2 AM UTC 触发的 CI 流程，包括矩阵生成、镜像导入和作业执行。
- 配置管理: `scripts/ci/slurm/nightly-configs.yaml` 以 YAML 格式声明基准测试配置，支持灵活扩展。
- 脚本执行: `launch_gb200.sh` 脚本克隆 `srt-slurm` 仓库、提交 Slurm 作业并监控日志；`generate_matrix.py` 动态生成 GitHub Actions 矩阵。
- 结果处理: `process_result.py` 解析原始 JSON 结果并添加拓扑元数据；`summarize.py` 汇总所有结果生成 Markdown 表格。

评论区精华

Review 讨论聚焦于设计权衡：

- 手动触发争议: Fridge003 提议添加 `workflow_dispatch` 以便测试，但作者 `csahithi` 回应：“I didn't add it initially to avoid anybody from launching jobs on the GB200”。最终，为安全起见，保持 `cron-only`。
- 失败处理优化: Fridge003 询问“Can we fail early when gb200 launching fails?”，`csahithi` 解释脚本已用 `set -e` 处理失败，并设置 360 分钟超时以避免资源抢占。
- 结果展示: Fridge003 建议“Maybe write some of them to github summary”，`csahithi` 确认已实现并提供了示例运行链接。

风险与影响

- 技术风险：脚本错误可能导致 CI 管道中断；配置错误可能使基准测试无效；依赖外部 srt-slurm 仓库变更可能破坏集成。
- 影响分析：对团队，自动化测试提升效率；对系统，新增夜间任务增加集群负载，但通过合理调度缓解；对用户，间接通过更稳定的性能保证受益。

关联脉络

与本 PR 相关的历史 PR 包括：

- PR 22465（更新 CI 权限）：同属 CI 基础设施改进，强化了权限管理。
- PR 21960（统一多模态 CI 测试）：展示了仓库在自动化测试 workflows 上的持续演进趋势。