

PR #23502 完整报告

sgl-project/sclang

[CI] Export GB200 nightly logs to S3

合并时间: 2026-04-24 06:35

原文链接: <http://prhub.com.cn/sgl-project/sclang/pull/23502>

执行摘要

- 一句话: 将 GB200 夜间测试日志导出到 S3 存储
- 推荐动作: 本 PR 属于基础设施增强, 逻辑清晰, 风险较低。值得关注的是其路径设计模式, 可复用于其他 CI 日志导出场景。建议合并。

功能与动机

目前查看 GB200 夜间运行日志的唯一方式是直接访问集群, 这给开发者排查失败问题带来了障碍。本 PR 将完整的 /logs 目录 (包括 prefill/decode/frontend worker 输出文件、benchmark-rollup.json 等) 上传到 S3 兼容后端, 方便开发者通过 `mc cp` 或 MinIO Web 控制台访问。

实现拆解

1. 修改 `scripts/ci/slurm/launch_gb200.sh`: 在生成的 `srtslurm.yaml` 中添加 `reporting.s3` 块, 包含 `bucket`、`prefix` 和 `endpoint_url`。`prefix` 由 `${TRIGGER}/${GITHUB_RUN_ID}-${GITHUB_RUN_ATTEMPT}/${SEQ_LEN}/${MATRIX_CONFIG_NAME}` 组成。同时添加了环境变量验证, 确保 S3 相关变量已设置。
2. 修改 `.github/workflows/nightly-72-gpu-gb200.yml`: 在 `nightly-gb200-benchmark` job 中添加 4 个环境变量 (`AWS_ACCESS_KEY_ID`、`AWS_SECRET_ACCESS_KEY`、`S3_BUCKET`、`S3_ENDPOINT_URL`), 从仓库 `secrets` 中获取。同时添加 `MATRIX_CONFIG_NAME` 变量。
3. 添加上传 AI 分析结果步骤: 在工作流中新增一个步骤, 当 job 失败时将 `ai_analysis.md` 文件上传到 S3, 路径与日志前缀一致, 便于关联分析。

关键文件:

- `scripts/ci/slurm/launch_gb200.sh` (模块 CI 脚本; 类别 `infra`; 类型 `infrastructure`): 核心变更: 添加 S3 配置块和环境变量验证, 定义日志路径前缀逻辑
- `.github/workflows/nightly-72-gpu-gb200.yml` (模块 CI 工作流; 类别 `infra`; 类型 `infrastructure`): 工作流配置: 添加 S3 环境变量和 AI 分析上传步骤

关键符号: `fmt_seq_len`

关键源码片段

scripts/ci/slurm/launch_gb200.sh

核心变更：添加 S3 配置块和环境变量验证，定义日志路径前缀逻辑

```
# 验证环境变量是否设置，若未设置则报错退出
: "${S3_BUCKET:?S3_BUCKET must be set}"
: "${S3_ENDPOINT_URL:?S3_ENDPOINT_URL must be set}"
: "${AWS_ACCESS_KEY_ID:?AWS_ACCESS_KEY_ID must be set}"
: "${AWS_SECRET_ACCESS_KEY:?AWS_SECRET_ACCESS_KEY must be set}"
: "${MATRIX_CONFIG_NAME:?MATRIX_CONFIG_NAME must be set}"
: "${GITHUB_RUN_ID:?GITHUB_RUN_ID must be set}"
: "${GITHUB_RUN_ATTEMPT:?GITHUB_RUN_ATTEMPT must be set}"

# 根据触发事件类型确定前缀顶级目录：cron 或 manual
case "${GITHUB_EVENT_NAME:-}" in
  schedule) TRIGGER=cron ;;
  workflow_dispatch) TRIGGER=manual ;;
  *) TRIGGER="${GITHUB_EVENT_NAME:-unknown}" ;;
esac

# 将 ISL/OSL 转换为 k 单位（如 1k1k），便于按序列长度分组
fmt_seq_len() {
  local n=$1
  if (( n % 1024 == 0 )); then echo "$((n / 1024))k"; else echo "$n"; fi
}
SEQ_LEN="$(fmt_seq_len "$ISL")$(fmt_seq_len "$OSL")"
S3_PREFIX="${TRIGGER}/${GITHUB_RUN_ID}-${GITHUB_RUN_ATTEMPT}/${SEQ_LEN}/${MATRIX_CONFIG_NAME}"

# 在 srtslurm.yaml 中添加 reporting.s3 配置，srt-slurm 会在 prefix 后追加日期 /Slurm 任务 ID
cat > srtslurm.yaml <<EOF
reporting:
  s3:
    bucket: "${S3_BUCKET}"
    prefix: "${S3_PREFIX}"
    endpoint_url: "${S3_ENDPOINT_URL}"
EOF
```

.github/workflows/nightly-72-gpu-gb200.yml

workflows配置：添加 S3 环境变量和 AI 分析上传步骤

```
# 在 job 的 env 中传递 S3 凭据，这些值来自 repo secrets
env:
  MATRIX_CONFIG_NAME: ${ matrix.config.name }
  AWS_ACCESS_KEY_ID: ${ secrets.NV_S3_ACCESS_KEY_ID }
  AWS_SECRET_ACCESS_KEY: ${ secrets.NV_S3_SECRET_ACCESS_KEY }
  S3_BUCKET: ${ secrets.NV_S3_BUCKET }
  S3_ENDPOINT_URL: ${ secrets.NV_S3_ENDPOINT_URL }

# 新增步骤：当 job 失败时，将 AI 分析结果上传到 S3
```

```
- name: Upload AI analysis to S3
  if: failure()
  continue-on-error: true
  env:
    ISL: ${ matrix.config.isl }
    OSL: ${ matrix.config.osl }
  run: |
    ANALYSIS="${ github.workspace }/ai_analysis.md"
    [ -f "$ANALYSIS" ] || { echo "no ai_analysis.md, skipping"; exit 0; }
    case "${ github.event_name }}" in
      schedule) TRIGGER=cron ;;
      workflow_dispatch) TRIGGER=manual ;;
      *) TRIGGER="${ github.event_name }}" ;;
    esac
    fmt() { if [ $(( $1 % 1024 )) -eq 0 ]; then echo "$(( $1 / 1024 ))k"; else echo "$1"; fi; }
    SEQ_LEN="$(fmt "$ISL")$(fmt "$OSL)"
    KEY="${TRIGGER}/${ github.run_id }-${ github.run_attempt }/${SEQ_LEN}/${ matrix.config.name }/ai_analysis.md"
    aws --endpoint-url "$S3_ENDPOINT_URL" s3 cp "$ANALYSIS" "s3://${S3_BUCKET}/${KEY}"
```

评论区精华

本 PR 仅有一条来自 gemini-code-assist bot 的评论，提示每日配额已达上限，未产生其他评审讨论。

- 暂无高价值评论线程

风险与影响

- 风险：
 1. 凭据泄露风险：AWS 凭据通过环境变量传递，未写入磁盘，但若 CI 日志泄露可能暴露凭据。建议确保 secrets 的访问权限受控。
 2. S3 端点可用性：依赖 MinIO 实例的稳定性，若端点不可用会导致上传失败（已设置 continue-on-error: true，不会阻塞主流程）。
 3. 路径冲突：多个运行可能产生相同前缀，可能导致覆盖。但 prefix 包含 run_id-attempt 和日期，冲突概率较低。- 影响：用户：开发者可通过 S3 访问日志，无需直接登录集群，降低排查故障的门槛。系统：新增 S3 上传步骤，使用 continue-on-error: true，不会影响核心测试流程。团队：需要维护 MinIO 实例和 S3 凭据，但无需变更现有 workflow。
- 风险标记：凭据通过环境变量传递，依赖外部 S3 端点

关联脉络

- 暂无明显关联 PR