

PR #28576 完整报告

sgl-project/sglang

[misc] Unify bench seed default to 42 and rename --profile-filename-prefix to --profile-prefix

合并时间: 2026-06-18 07:41

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/28576>

执行摘要

- 一句话: 统一 bench seed 默认值并重命名 profile 参数
- 推荐动作: 本 PR 为低风险的机械性清理, 可直接通过。若团队有依赖默认 seed=1 的自动化基准测试, 需注意结果可能微变, 但通常 42 更规范。建议在后续 PR 中进一步统一其他参数名称 (如 result-filename 等)。

功能与动机

PR body 指出这是 "Mechanical consistency cleanup across the bench scripts", 旨在统一不同 bench 脚本之间的参数默认值和命名风格, 降低用户混淆和维护成本。具体来说, `bench_serving` 和 `bench_offline_throughput` 的 `--seed` 默认值为 1, 而 `bench_one_batch_server` 已使用 42; `bench_one_batch` 的 `--profile-filename-prefix` 与其他脚本中的 `--profile-prefix` 不统一。

实现拆解

1. 统一 `--seed` 默认值: 在 `bench_offline_throughput.py` 和 `bench_serving.py` 中, 将 `BenchArgs.seed` 字段默认值和对应的 `argparse add_argument` 调用的 `default` 参数从 1 改为 42。
2. 重命名 `--profile-filename-prefix` 为 `--profile-prefix`: 在 `bench_one_batch.py` 中:
 - 将 `BenchArgs` 数据类的字段名从 `profile_filename_prefix` 改为 `profile_prefix`。
 - 在 `add_cli_args` 中将 `--profile-filename-prefix` 替换为 `--profile-prefix`, 同时将旧参数名作为别名保留, 通过 `dest="profile_prefix"` 解析到同一目标。
3. 更新下游引用: 在 `bench_one_batch.py` 中, 将 `_create_torch_profiler_filename` 函数签名及其调用处、以及 `latency_test` 和 `latency_test_run_once` 函数签名中的参数名同步修改为 `profile_prefix`。
4. 更新帮助字符串: 相应更新了参数帮助文本中的占位符。

关键文件:

- `python/sglang/bench_one_batch.py` (模块 基准测试; 类别 source; 类型 core-logic; 符号 `BenchArgs`, `add_cli_args`, `latency_test`, `latency_test_run_once`): 重命名 `profile_filename_prefix` 为 `profile_prefix`, 涉及数据类字段、`argparse` 参数声明、函数签名及调用链的多处修改, 并保留旧参数作为别名。

- python/sclang/bench_offline_throughput.py (模块 基准测试; 类别 source; 类型 core-logic; 符号 BenchArgs, add_cli_args) : 将 --seed 默认值从 1 改为 42, 涉及数据类字段和 argparse 参数两处。
- python/sclang/bench_serving.py (模块 基准测试; 类别 source; 类型 core-logic; 符号 add_cli_args) : 将 --seed 默认值从 1 改为 42, 仅 argparse 参数处一处修改。

关键符号: 未识别

关键源码片段

python/sclang/bench_one_batch.py

重命名 `profile_filename_prefix` 为 `profile_prefix`, 涉及数据类字段、argparse 参数声明、函数签名及调用链的多处修改, 并保留旧参数作为别名。

bench_one_batch.py 中关键变更: 参数名统一与别名兼容

```
@dataclasses.dataclass
class BenchArgs:
    # ... 其他字段 ...
    profile_prefix: str = "profile" # 原为 profile_filename_prefix
    # ...

    @staticmethod
    def add_cli_args(parser: argparse.ArgumentParser):
        # ... 其他参数 ...
        parser.add_argument(
            "--profile-prefix", # 新名称作为主参数
            "--profile-filename-prefix", # 旧名称作为 deprecated alias
            dest="profile_prefix", # 解析到同一字段
            type=str,
            default=BenchArgs.profile_prefix,
            help='Prefix of the profiling file names. ...',
        )
        # ...

    def _create_torch_profiler_filename(
        profile_prefix, batch_size, input_len, output_len, stage # 参数名更新
    ):
        output_dir = _get_torch_profiler_output_dir()
        # 占位符同步更新
        filename = f"{profile_prefix}_batch{batch_size}_input{input_len}_output{output_len}_{stage}.
        trace.json.gz"
        return os.path.join(output_dir, filename)
```

评论区精华

PR 无 review 讨论, 无 comment。作者独立完成并合并。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低。变更仅涉及参数默认值和参数名称；旧参数名 `--profile-filename-prefix` 通过 `alias` 保留，不会破坏现有调用脚本。默认种子变更可能影响依赖固定随机性的基准测试结果，但 42 是更常用的标准种子，且影响可预见。无核心逻辑改动，未涉及测试文件变更（但 PR body 未声明需要测试配套）。
- 影响：
 - 用户影响：使用默认 `seed` 的基准测试结果可能略有变化（因随机性），但无关紧要；使用 `--profile-filename-prefix` 的用户会收到 `deprecation` 警告（`argparse` 自动处理别名），但脚本仍能正常工作。
 - 系统影响：无。
 - 团队影响：降低了参数不一致带来的维护负担，提升了代码可读性。
 - 风险标记：无风险，纯机械性清理

关联脉络

- PR #28563 [Fix] Reuse an already-running server in `bench_one_batch_server` instead of forking an orphan: 同样涉及 `bench_one_batch` 参数和功能，且该 PR 修改了 `bench_one_batch_server_internal.py`，本 PR 重命名了 `bench_one_batch` 中的参数，可能存在关联。
- PR #28483 [Fix] DeepSeek-OCR-2 `bench_serving`: fix processor loading: 涉及 `bench_serving` 的修复，与本 PR 同属基准测试脚本维护方向。
- PR #22053 [HiCache & Bench] add cache hit breakdown in `bench_serving`: 上一轮 `bench_serving` 功能增强，本 PR 进一步统一了 `seed` 默认值。