

# PR #23355 完整报告

sgl-project/sglang

Remove smoke wording from tests and comments

合并时间: 2026-04-28 12:05

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/23355>

## 执行摘要

- 一句话: 移除测试和注释中的 'smoke' 措辞
- 推荐动作: 该 PR 属于纯代码清理, 不涉及功能变更, 适合在代码审查中快速过目。对于追求代码术语一致性的项目, 这是一个可取的实践。建议关注 `gemini-code-assist` 提出的不一致问题, 确保后续修复。

## 功能与动机

从 PR body 可见, 目的是替换 'smoke' 措辞为更具体的标签 (如 `sanity check`, `build-and-run check`, `validation settings`), 并重命名 CUDA coredump debug 测试文件 / 类避免 'smoke' 标签, 同时保持 `benchmark/test prompt` 载荷不变。

## 实现拆解

1. 全局搜索与替换: 使用 `git grep` 定位所有包含 'smoke' 的测试名、类名和注释。
2. 测试函数重命名: 将 `test_generate_smoke` 改为 `test_generate_basic`, `test_generate_smoke_tp_mismatch` 改为 `test_generate_with_tp_mismatch`, `test_text_output_smoke` 改为 `test_text_output_format`。
3. 测试文件重命名: 将 `test_cuda_coredump_smoke.py` 重命名为 `test_cuda_coredump.py`, 类名 `TestCudaCoredumpSmoke` 改为 `TestCudaCoredump`。
4. 注释措辞调整: 将 `benchmark` 中的 `'# Smoke test'` 改为 `'# Validate output shapes before benchmarking.'`; `deepseek-v4-deployment.jsx` 中的 `'static smoke'` 改为 `'static validation'`; `diffusion trajectory` 工具中的 `'smoke settings'` 改为 `'deterministic settings'`。
5. 配套修改: 更新了多处的 Skill 文档、CI workflow 和测试配置文件中的 'smoke' 相关措辞。
6. 验证: 通过 `git diff --check` 和 `git grep -n -I -i 'smoke'` 确认无残留。

关键文件:

- `test/registered/debug_utils/test_cuda_coredump.py` (模块 `CUDA` 调试; 类别 `test`; 类型 `rename-or-move`; 符号 `TestCudaCoredumpSmoke`, `TestCudaCoredump`): 核心变更: 测试文件重命名并移除类名中 'Smoke', 体现 PR 主要意图。
- `test/manual/test_mori_transfer_engine_e2e.py` (模块 `传输引擎`; 类别 `test`; 类型 `test-coverage`; 符号 `test_generate_smoke`, `test_generate_basic`,

test\_generate\_smoke\_tp\_mismatch, test\_generate\_with\_tp\_mismatch) : 涉及两个测试函数重命名, 展示 'Smoke' 替换为 'Basic'、'With' 等。

- benchmark/bench\_rope/benchmark\_rope\_index.py (模块 基准测试; 类别 source; 类型 core-logic) : 注释清理示例: '# Smoke test' 改为 '# Validate output shapes before benchmarking.'
- docs\_new/src/snippets/autoregressive/deepseek-v4-deployment.jsx (模块 文档片段; 类别 source; 类型 core-logic) : 文档示例中 'static smoke' 改为 'static validation'
- python/sglang/multimodal\_gen/tools/compare\_diffusion\_trajectory\_similarity.py (模块 多模态生成; 类别 source; 类型 core-logic) : 注释清理: 'smoke settings' 改为 'deterministic settings'
- test/registered/debug\_utils/comparator/test\_entrypoint.py (模块 入口测试; 类别 test; 类型 test-coverage; 符号 test\_text\_output\_smoke, test\_text\_output\_format) : 测试方法重命名: test\_text\_output\_smoke→test\_text\_output\_format
- python/sglang/test/attention/test\_trtllm\_mla\_backend.py (模块 MLA 后端; 类别 test; 类型 test-coverage) : 文本清理, 涉及 'comprehensive' 等措辞调整
- test/registered/hicache/test\_hicache\_storage\_runtime\_attach\_detach.py (模块 HiCache 存储; 类别 test; 类型 test-coverage) : 附带清理, 重要性较低
- test/registered/moe/test\_cuteds1\_moe.py (模块 MoE 测试; 类别 test; 类型 test-coverage) : 附带清理, 重要性较低
- .github/workflows/pr-test-ssl-kernel.yml (模块 CI 配置; 类别 infra; 类型 infrastructure) : CI 工作流程中 'comprehensive' 等措辞调整
- python/sglang/multimodal\_gen/.claude/skills/sglang-diffusion-modelopt-quant/SKILL.md (模块 多模态文档; 类别 docs; 类型 documentation) : 文档清理, 涉及控制流 / 异常路径术语调整
- .claude/skills/sglang-auto-benchmark/SKILL.md (模块 Claude 技能; 类别 docs; 类型 documentation) : 文档清理, 重要性最低

关键符号: TestCudaCoredump, test\_generate\_basic, test\_generate\_with\_tp\_mismatch, test\_text\_output\_format

## 关键源码片段

### test/manual/test\_mori\_transfer\_engine\_e2e.py

涉及两个测试函数重命名, 展示 'Smoke' 替换为 'Basic'、'With' 等。

```
def test_generate_basic(self): # 原名为 test_generate_smoke
    resp = requests.post(
        self.lb_url + "/generate",
        json={
            "text": "Hello",
            "sampling_params": {"temperature": 0, "max_new_tokens": 8},
        },
        timeout=120,
    )
```

```
self.assertEqual(resp.status_code, 200, resp.text)
out = resp.json()
self.assertIn("text", out)
self.assertIsInstance(out["text"], str)
self.assertGreater(len(out["text"]), 0)
```

# ... 另一处类似:

```
def test_generate_with_tp_mismatch(self): # 原名为 test_generate_smoke_tp_mismatch
    # 测试体相同
```

## 评论区精华

gemini-code-assist[bot] 对 [test/manual/layers/attention/nsa/test\\_act\\_quant\\_triton.py](#) 提出两个 medium 优先级的评论: 更新后的 docstring 提到 'with and without CUDA graphs', 但对应的 print 语句只提到 'with CUDA Graphs', 建议保持一致。PR 作者未公开回复, 但从 PR 合并状态看该问题未在当前 PR 中修复 (该文件未在变更集中)。

- test\_act\_quant\_triton.py docstring 和 print 不一致 (testing): 建议确保一致性, 但该文件未在当前 PR 的变更集中, 状态未知。

## 风险与影响

- 风险: 风险极低: 所有修改均为注释、测试命名和类名, 不涉及运行时逻辑。潜在风险是测试重命名可能导致外部脚本或 CI 配置引用旧的测试名称, 但 PR 验证了 git grep 无 'smoke' 残留, 且项目内无其他引用。
- 影响:
  - 用户: 无感知, 不影响任何功能。
  - 系统: 无性能变化, 无行为变化。
  - 团队: 改善了代码可读性, 减少模糊术语。但测试名称变更可能需要团队成员更新个人工作流中的测试选择器。
  - 风险标记: 仅注释和命名变更, 无逻辑影响

## 关联脉络

- 暂无明显关联 PR