

PR #23712 完整报告

sgl-project/sglang

[Doc]Add msprobe doc in docs_new path

合并时间: 2026-04-27 15:06

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/23712>

执行摘要

- 一句话: 在 docs_new 中添加 MSProbe 调试指南
- 推荐动作: 建议快速合并。此 PR 为团队开发者提供了重要的调试工具文档, 评审已通过, 技术争议已澄清。对于关注 NPU 调试和模型精度问题的工程师, 值得仔细阅读文档中的配置与示例部分。

功能与动机

根据 PR 描述, 目的是在 docs_new 路径中添加 msprobe 文档, 为开发者提供 MSProbe 调试工具的使用指导。

实现拆解

1. 创建核心文档: 新增 docs_new/docs/developer_guide/msprobe_debugging_guide.mdx, 包含 MSProbe 的背景概念 (L0/L1/mix 转储级别)、安装方法、关键参数配置表 (task、dump_path、rank、step、summary_mode、list 等), 以及端到端调试示例与故障排查步骤。
2. 注册导航路径: 修改 docs_new/docs.json, 在 Developer Guide 的 pages 数组末尾添加 "docs/developer_guide/msprobe_debugging_guide", 使文档在侧边导航中可访问。
3. 更新概览页面: 在 docs_new/docs/developer_guide/overview.mdx 的连接列表末尾追加一行 - [MSProbe Debugging Guide](./msprobe_debugging_guide), 确保用户可从开发指南总览页跳转。
4. 无测试或配置配套变更: 本次为纯文档添加, 未涉及代码、测试或部署配置。

关键文件:

- docs_new/docs/developer_guide/msprobe_debugging_guide.mdx (模块 开发者文档; 类别 other; 类型 documentation): 新增的核心文档, 完整介绍 MSProbe 工具的使用方法、参数配置、端到端示例和故障排查。是本次 PR 的主要产出。
- docs_new/docs.json (模块 站点配置; 类别 config; 类型 configuration): 站点导航配置文件, 新增文档条目使其出现在侧边栏 Developer Guide 组中。
- docs_new/docs/developer_guide/overview.mdx (模块 开发者文档; 类别 other; 类型 documentation): 开发指南概览页, 新增链接便于用户跳转到新的调试指南。

关键符号: 未识别

评论区精华

Review 由 gemini-code-assist[bot] 发起，共提出 4 条中等优先级评论：

- 术语“pynative”：建议替换为 PyTorch 生态的“eager mode”，以避免混淆（ping1jing2 赞同）。
- 格式空格：建议在括号前添加空格以保持排版一致性。
- 拼写错误：srep0 应为 step0。
- 技术不一致：summary_mode="md5" 时描述中却提及 CRC-32，is-not 回应“这是命名问题，实际转储为 CRC-32”。

上述评论除 CRC-32 命名获作者澄清外，其余预计已在后续提交中修正。

- 术语“pynative”应替换为“eager mode” (documentation): Reviewer 建议被采纳，预计文档已修改为使用“eager mode”。
- 格式：括号前缺少空格 (style): 格式问题已根据建议修正。
- 技术不一致：CRC-32 与 MD5 的混淆 (correctness): 作者澄清命名一致，文档无需修改，争议解决。

风险与影响

- 风险：文档变更不涉及运行代码，无回归、性能或安全风险。主要风险在于术语准确性：若保留“pynative”等 MindSpore 特定术语，可能对 PyTorch 用户造成困惑；但作者已在评论中澄清 CRC-32 的命名逻辑，整体影响可控。
- 影响：对开发者而言，新增了一篇完整的调试指南，尤其对使用 NPU 进行模型精度调试的场景帮助明显。对系统无影响。文档变更范围局限于 docs_new 目录，不影响任何运行时行为。
- 风险标记：术语准确性，文档内容待完善

关联脉络

- 暂无明显关联 PR