

PR #7452 完整报告

verl-project/verl

[doc] chore: Update ckpt engine readme

合并时间: 2026-08-18 10:55

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7452>

执行摘要

- 一句话: Ascend 950 文档更新: kimi checkpoint engine 改白名单支持
- 推荐动作: 可以花少量时间精读这段文档, 重点是理解 HCCL 白名单机制在 Ascend 950 上的启用方式; 建议在合并后跟进修正 JSON 示例格式, 并补充 IP 格式与适用范围说明。若后续有相关源码支持 PR, 可与此文档同步更新。

功能与动机

PR body 明确目标是 `add kimi-ckpt-engine support in ascend 950`, 说明此前 Ascend 950 被认为不支持 `kimi-checkpoint-engine`, 实际上可通过 HCCL 白名单机制开启。README 修改正是为了纠正这一认知偏差, 并给出可操作的环境变量配置方法。

实现拆解

1. 变更入口: 修改 `verl/checkpoint_engine/README.md` 的“Important Configuration for Ascend Devices”小节。
2. 核心内容: 将“`kimi-checkpoint-engine hasn't been supported in Ascend 950`”改为需设置环境变量 `HCCL_WHITELIST_DISABLE=0` 与 `HCCL_WHITELIST_FILE=/path/to/whitelist.json`, 并创建 JSON 白名单文件登记设备 IP。
3. 示例补充: 给出 whitelist JSON 文件的格式模板; 原 PR 中的写法 `{ "host_ip": ["ip1"], ["ip2"] }` 不是合法键值对, 本报告按合理语义整理为 `{ "host_ip": ["ip1", "ip2"] }`。
4. 兼容衔接: 保留原有“CANN >= 8.5.0 时需设置环境变量启用 node 内 ROCE”的提示, 确保新旧配置说明连续。
5. 配套说明: 无源码、测试、配置模板改动, 纯文档变更。

关键文件:

- `verl/checkpoint_engine/README.md` (模块 检查点引擎; 类别 docs; 类型 documentation): 唯一变更文件, 说明 Ascend 950 上 `kimi-checkpoint-engine` 的白名单配置方式, 是本次 PR 的核心交付。

关键符号: 未识别

评论区精华

该 PR 没有实质 review 评论；唯一的审核人 [wucong25](#) 直接以 **APPROVED** 状态通过，未留下讨论或反对意见。

- 暂无高价值评论线程

风险与影响

- 风险：主要风险在文档正确性：README 中的 JSON 示例 { "host_ip": ["ip1"], ["ip2"] } 不是合法键值对，["ip2"] 缺少键，用户照抄可能导致配置失败；按语义应为 { "host_ip": ["ip1", "ip2"] }。此外，文档未说明 HCCL_WHITELIST_FILE 的路径要求、IP 格式（IPv4/IPv6）以及是否适用于其他 Ascend 型号，信息完整性不足。由于是纯文档变更，无运行时影响，技术风险极低。
- 影响：影响范围为用户使用 Ascend 950 并尝试启用 kimi-checkpoint-engine 的用户：本次文档使其从“认为不支持”变为可配置，降低上手门槛。对仓库代码功能和现有测试无影响。对团队而言，文档维护成本低，但 JSON 示例若被照搬可能导致用户配置报错，需关注后续反馈或补丁。
- 风险标记：JSON 示例格式存疑，信息完整性不足

关联脉络

- PR #7291 [ckpt] feat: Node-local multi-sender broadcast in NCCL checkpoint engine: 同属 checkpoint_engine 模块，本期 README 补充了 kimi-checkpoint-engine 在 Ascend 950 上的使用说明，与该 PR 的 checkpoint engine 能力演进相关。