

# PR #1749 完整报告

THUDM/slime

Update MTP example configs, rename GLM-4.5 to GLM-4.7, clean scripts

合并时间: 2026-03-22 16:24

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1749>

## 执行摘要

本次 PR 主要更新了 MTP 示例配置，将 GLM-4.5 重命名为 GLM-4.7，并清理了相关脚本和文档。这是一个常规维护性变更，旨在同步最新模型版本和优化示例配置，对用户影响较小但需注意文档引用更新。

## 功能与动机

为什么做: 根据 PR body 描述，目的是重命名 GLM-4.5 为 GLM-4.7，以反映模型升级；同时同步已验证的配置（如 Qwen3.5-27B 的 4 节点 32k GRPO 配置）并清理冗余脚本，提升示例的准确性和简洁性。

## 实现拆解

按模块拆解改动:

- 文档模块: 删除旧版 GLM-4.5 文档（如 docs/en/examples/glm4.5-355B-A32B.md），新增 GLM-4.7 文档（如 docs/en/examples/glm4.7-355B-A32B.md），并更新索引和快速入门指南中的链接。
- 脚本模块: 重命名运行脚本（如 scripts/run-glm4.5-355B-A32B.sh 改为 run-glm4.7-355B-A32B.sh），修改配置参数，例如：
  - 在 scripts/run-glm4.7-30B-A3B.sh 中，将 TP 从 4 改为 2，PP 从 1 改为 2，并更新投机解码步骤数。
  - 在 scripts/run-deepseek-r1.sh 中添加 EAGLE MTP 投机解码参数。
- 清理工作: 移除内部集群环境变量（如 NCCL\_IB\_\*），删除冗余脚本 scripts/run-qwen3.5-35B-A3B.sh。

## 评论区精华

讨论摘要: review 评论为空，表明本次变更未经过讨论，直接合并执行。

## 风险与影响

具体风险:

- 文档风险: 旧文档链接失效，用户需更新引用；新增文档可能未充分测试配置。

- 兼容性风险：脚本参数更改（如并行设置）可能在某些部署环境中引发问题，建议验证新配置。
- 简化风险：清理环境变量可能影响特定集群的网络配置。

影响评估：

- 用户：需要适应新文档和脚本，但配置优化可能提升训练效率。
- 系统：文档和脚本更清晰，减少维护负担。
- 团队：支持 GLM-4.7 模型，保持示例库的时效性。

## 关联脉络

与历史 PR 的关系：

- 近期 PR 如 #1721（添加 Qwen3.5-4B 支持）和 #1742（优化 qwen3.5 loss mask）同样涉及模型配置更新，表明仓库正在持续维护和扩展模型示例库。
- 本 PR 是这一趋势的一部分，专注于 GLM 模型版本的更新和配置同步，未引入新功能但强化了文档和脚本的可用性。