

# PR #5491 完整报告

verl-project/verl

[doc] feat: add mtp spec log

合并时间: 2026-03-05 11:06

原文链接: <http://prhub.com.cn/verl-project/verl/pull/5491>

## 执行摘要

本次 PR 在 MTP 文档中添加了性能说明和图像，记录启用 MTP 对 rollout 接受率和 H20 GPU吞吐量的影响，变更简单但暴露了外部图像托管的风险，建议内部化资产以确保长期稳定。

## 功能与动机

本 PR 旨在通过文档更新，记录 MTP 启用后的性能数据：启用 MTP 可以提升 rollout 接受率约 14%，但在 H20 GPUs 上整体吞吐量未增加甚至轻微下降。动机来自 PR body 中用户需要了解 MTP 在不同硬件上的性能权衡，以做出更明智的部署决策。

## 实现拆解

变更仅涉及一个文件 `docs/advance/mtp.md`，在 'Performance Notes for MTP in Rollout Inference' 章节下添加了以下内容：

- 文本描述: "Enabling MTP improves the rollout acceptance rate by around 14%. However, on H20 GPUs, overall throughput does not increase and even decreases slightly."
- 图像链接: `![spec_log](https://github.com/ArronHZG/verl-community/blob/main/docs/spec_log.png?raw=true)` 无代码逻辑变更，实现简单直接。

## 评论区精华

review 中 `gemini-code-assist[bot]` 评论指出关键问题：

"This image is hosted on a personal GitHub repository. To ensure the long-term availability and maintainability of the documentation, it's best practice to store all assets, like images, within the same repository." 此讨论聚焦于文档维护性，建议内部化图像资产以避免未来链接失效，但评论未被处理，PR 直接合并。

## 风险与影响

- 风险: 图像托管在外部个人仓库，如果仓库变更或删除，将导致文档图像失效，影响用户体验。无其他技术风险。
- 影响: 仅影响文档层面，提升用户对 MTP 性能的理解，但对系统功能和性能无直接影响；长期看，外部依赖可能增加维护负担。

## 关联脉络

与近期历史 PR 5763（同为文档更新）关联，显示团队持续优化文档以支持硬件性能说明。这反映了 verl 项目中对文档维护的重视，但外部资产托管问题可能成为跨 PR 的共同风险点。