

verl 周报：2026 第 33 周（08-10 至 08-16）

verl-project/verl

周期：2026-08-10 至 2026-08-16

来源 PR：33 · 重点 PR：24 · 自动生成

原文链接：<http://prhub.com.cn/verl-project/verl/reports/2026-08-10-to-2026-08-16>

verl 2026 第 33 周周报（08-10 至 08-16）

执行摘要

本周 verl 共合入 33 个 PR，其中 24 个为重点 PR，整体呈现三条主线：一是 Megatron 训练路径的正确性与性能收敛，围绕 router replay、动态上下文并行（DCP）、int16 路由索引与 DeepSeek-V4 CP 布局等多点修复；二是 NPU/Ascend CI 的稳定性加固与基础环境升级，nightly_ascend.yml 等文件被反复修改；三是 LoRA 同步链路（SGLang/vLLM）端到端打通，连续多个 PR 修复权重名解析与适配器同步缺陷。与此同时，mindspeedllm 后端移除、Continuous Token 成为唯一 tokenization 路径等破坏性变更，表明架构正加速向统一和收敛的方向演进。

本周重点变化

Megatron 链路密集修复。Qwen3.5 router replay 静默失效（#7340）源于进程级全局列表在复杂构建流程下寻址错误，修复后改为按实际模型对象树寻址；DeepSeek-V4 的 CP>1 路径（#7297）通过透传 `cp_partition_mode` 打通；mHC+MTP 训练崩溃（#7328）则定位到激活回收与多流张量转发的两处缺陷。这些 PR 共同补齐了关键模型的训练能力。

性能优化显著。动态 CP 调度（#6555）以 thin adapter 复用 Megatron-Core 的 `DefaultDynamicCPScheduler`，在长尾序列分布下吞吐提升约 60%；NCCL checkpoint 多发送者广播（#7291）让同节点权重同步提速约 3 倍；路由索引升级到 int16（#7407）则兼顾了超过 255 专家的正确性与内存带宽。

LoRA 同步链路打通。#7413 修复了 LoRA 从训练引擎同步到 SGLang 的 7 处链式缺陷，#7327 与 #7376 分别解决 vLLM 接收端的 `.base_layer` 解析与混合 LoRA 模型的权重名归一化，使该组合首次端到端跑通。

架构清理与破坏性变更。#7374 删除 mindspeedllm 后端约 1941 行；#6804 将 Continuous Token 扩展为多模态默认路径并移除开关；#7357 将 CI 工作流迁移到 v1 separate_async。这些变更意味着旧路径不再维护，用户需要跟随迁移。

模块与主题趋势

Megatron 是绝对焦点，13 个 PR 带 megatron 标签，hot_files 中 `verl/workers/engine/megatron/transformer_impl.py` 被改 5 次、`verl/utils/megatron/router_replay_utils.py` 被改 4 次，说明引擎层与路由回放是当前核心攻坚点。NPU/Ascend 是第二大主题（10 个 npu、7 个 ci），团队在昇腾环境上投入大量精力

修复依赖、镜像与 nightly 任务。权重同步 /LoRA 则形成一个小集群，涉及 vllm、sglang、peft 等多个标签，体现了多后端一致性维护的复杂度。此外，文档、算法与 Ray 基础设置也有分散更新，但整体上工程化清理（版本 bump、CI 迁移、后端移除）占了相当比例。

风险观察

本周最大的风险是测试覆盖不足。“缺少测试覆盖”在风险统计中出现 6 次，多个核心路径 PR（如 LoRA 同步、NCCL checkpoint、动态 CP）没有新增单测，依赖端到端验证，回归风险较高。其次，多项 BREAKING 变更集中落地，用户自定义配置和工作流需要同步调整。int16 路由索引属于核心数据契约跨模块变更，且 NCCL 不支持 int16，依赖位视图绕行，后续需关注与其他模块的兼容。此外，多个特性依赖较新的 Megatron-Core 或 sglang 版本，版本锁定与升级策略需要明确。最后，NCCL checkpoint 的 multi_sender 实际默认开启但文档仍写默认关闭，属于文档与实现不一致的隐患。

重点 PR 速览

PR	标题摘要	重要性	核心要点
#6804	Multimodal Continuous Token	9.4	BREAKING: CT 成为唯一 tokenization 路径，新增多个 VL builder
#6555	动态 CP 调度	9.4	复用 Megatron-Core 调度器，长尾序列吞吐提升约 60%
#7407	int16 routed_experts	9.2	路由索引升级，支持 >255 专家，NCCL 位视图绕行
#7374	移除 mindspeedllm 后端	9.0	删除约 1941 行 legacy 代码，NPU Megatron 仍保留
#7027	reward 路径确定性	8.6	补齐全确定性训练，生成式 RM 按位可复现
#7327	vLLM 非合并 LoRA 解析	8.7	接收端按命名空间解析 .base_layer，覆盖视觉 / 共享专家场景
#7291	NCCL 多发送者广播	8.4	权重同步提速约 3 倍，用 Ray 节点 id 推断 NVLink 可达性
#7340	Router replay 静默失效	8.4	从“列表位置”改为“对象树”寻址，修复 Qwen3.5
#7413	SGLang LoRA e2e 修复	7.9	一次性修复 7 处链式缺陷，端到端首次跑通
#7300	REINFORCE++ 奖励修复	7.1	observation span 不再清零 running_return，与 GAE 行为对齐

这些 PR 之间存在明显关联：#7407 与 #7340 都落在 router replay 链路上，#7327/#7376/#7413 共同支撑 LoRA 同步，#7291/#7107 则是 checkpoint 性能组合。它们大多由少量维护者评审（wuxibin89 出现频率最高），讨论虽少但关键设计交锋质量较高。

后续建议

1. 补测试：为 LoRA 同步、NCCL checkpoint 多发送者、动态 CP 等核心路径补充单元测试或 CI 集成，避免回归。
2. 跟进遗留缺口：Qwen3.5 VL 的 family 注册缺失、multi_sender 文档不一致应尽快修掉。
3. 版本锁定：将 Megatron-Core、MindSpeed、vLLM、sglang 等依赖版本集中到锁文件或常量，降低漂移风险。
4. 迁移指南：针对 BREAKING 变更（CT 路径、mindspeedllm 移除、CI 迁移）撰写升级说明，帮助用户平滑过渡。
5. 推广基线对比：NPU nightly 的基线对比模式可推广到 GPU nightly，提前发现性能回退。