

verl-project/verl 2026 年第 25 周 (06-15 至 06-21) 周报

verl-project/verl

周期: 2026-06-15 至 2026-06-21

来源 PR: 21 • 重点 PR: 21 • 自动生成

原文链接: <http://prhub.com.cn/verl-project/verl/reports/2026-06-15-to-2026-06-21>

执行摘要

本周 (2026 年 6 月 15 日 -6 月 21 日) verl 仓库共合并 21 个 PR, 平均重要性 6.12, 平均洞察 4.48。重点集中在训练流程拆解、显存优化、多后端支持以及基础设施升级。其中, Tinker 训练原语、FSDP+LoRA 分层召唤、Megatron 分块熵和 SGLang 权重同步优化是本周亮点。同时, 测试覆盖不足和部分配置风险值得关注。团队在本周表现出较高的交付效率, 但需要加强对遗留风险的跟踪。

本周重点变化

1. 训练流程拆解: TinkerTrainingWorker (#6717) 将梯度清零、前反向传播、优化器步骤暴露为独立 RPC, 实现显式梯度累积控制; 同步引入优化器参数覆盖 (#6765) 允许动态调整学习率等超参。分离异步 Trainer (#6790) 从 stub 变为可运行, 为大规模异步训练提供基础。
2. 显存优化路径: FSDP+LoRA 分层召唤 (#6512) 通过逐个 FSDP 单元 all-gather 替代全模型 all-gather, 显著降低峰值显存, 并修复 wrap 策略冲突导致的 NCCL 死锁, 但嵌套 FSDP 单元重复收集问题尚未彻底解决; Megatron 分块熵 (#6446) 在张量并行环境下沿序列维度分块计算熵, 有效降低长序列显存占用; SGLang 权重同步 (#6738) 通过 `_compact_for_bucket` 函数避免冗余张量 clone, 消除多 GiB 瞬态内存。
3. 离线训练能力增强: PPO ReplayBuffer 新增陈旧性控制策略 (#6778) 支持 'wait' 和 'drop' 两种模式, 并引入 off-policy 指标 (#6736) 监控策略陈旧程度。但指标未过滤 padding 序列, 计算精度受影响。
4. 多后端与硬件支持扩展: ROCm Docker 重构 (#6619) 引入 ccache 和 BuildKit 缓存加快构建; SGLang ROCm 后端 (#6664) 通过平台抽象层注入环境变量, 自动选择 aiter attention 后端; Megatron Lite 示例脚本 (#6791) 覆盖 DeepSeek-V4、Kimi K2.6、GLM 5.1 等模型; Ascend 双节点 CI (#6672) 上线, 保障 NPU 多机训练稳定性。
5. 基础设施与依赖升级: NumPy 约束放宽至 `>=2.0.0` (#6551) 解决包管理冲突, 但可能影响 NPU 路径; vLLM 与 NCCL 同步升级, 并补丁 vLLM 未合并修复; TRTLLM 镜像升级至 1.3.0rc15 (#6730) 但使用了不存在的 `transformers==5.3.0`, 存在构建风险。

模块与主题趋势

- Trainer 模块最为活跃, 涉及 Tinker 原语、分离异步支持、replay buffer 策略扩展、指标添加, 表明团队正加速训练流程的可编程化和异步化。

- Rollout 模块聚焦显存优化和硬件兼容，SGLang OOM 修复和 ROCm 后端适配减轻了特定场景的内存压力，vLLM 配置兼容性补丁维护了向后兼容。
- Infra/Docker/CI方面，依赖升级和镜像重构提升了构建可重复性和环境一致性，但版本错误和参数传递问题需警惕。
- 模型与 NPU方面，NPU 补丁兼容 transformers 5.x (PR #6708)，保障了 Qwen3 MoE 在多版本下的稳定运行。
- 文档与测试：OPD 文档修复、Ascend 安装指南改进；测试覆盖在部分 PR 中得到增强，但整体仍显不足。

风险观察

- FSDP+LoRA 重复收集风险(#6512): PR 中的 layered_summon_lora_params 遍历 FSDP 单元时可能重复收集子单元参数，官方回复“fixed”但代码未显式去重，需在集成验证中重点关注。
- NumPy 2.0 与 NPU 兼容(#6551): NPU 路径当前通过 uv 固定 numpy<2.0，未来启用 2.0 时需全面测试 torch-npu 兼容性。
- TRTLLM 构建失败(#6730): transformers==5.3.0 在 PyPI 不存在，Docker 构建将直接失败，需紧急修复。
- Off-policy 指标准确性(#6736): 指标计算未过滤 padding，监控数据有效性存疑。
- 测试覆盖缺口：多个关键 PR (#6446、#6664、#6708) 未附带测试，回归风险高。

重点 PR 速览

PR #	标题	模块	重要性	关键点
#6717	Tinker 训练 Worker 原语	trainer	9.1	暴露 optimizer_zero_grad/forward_backward/optimizer_step 为独立 RPC，5 个后端支持
#6512	FSDP+LoRA 分层召唤	fsdp	7.8	逐个 FSDP 单元 all-gather 降低显存，修复死锁，但嵌套重复收集风险
#6778	ReplayBuffer 陈旧性控制	trainer	8.6	支持 drop/wait 策略，优化离线采样
#6446	Megatron 分块熵	megatron	7.0	序列分块降低长训练显存，配置开关

PR #	标题	模块	重要性	关键点
#6738	SGLang 权重同步 OOM 修复	rollout	7.0	消除冗余 clone, 节省多 GiB 显存
#6551	依赖升级 NumPy /vLLM/NCCL	infra	5.9	解决包冲突, 升级核心依赖
#6790	分离异步 Trainer 可运行	trainer	6.7	从 stub 到可运行, 关键一步
#6619	ROCm Docker 重构	docker	5.3	ccache 缓存、BuildKit 挂载, 构建提速

后续建议

- 修复 TRTLLM 版本问题: 立即将 transformers==5.3.0 改为有效版本 (如 4.43.0), 避免 CI 阻塞。
- 完善测试覆盖: 为 Megatron 分块熵、SGLang ROCm 后端、NPU 补丁等添加单元测试和集成测试, 降低回归风险。
- 跟踪 FSDP+LoRA 风险: 对 #6512 进行针对性的长序列显存测试, 确认重复收集问题已彻底解决, 必要时推动作者补充显式去重。
- 优化 Off-policy 指标: 采纳 review 建议, 在指标计算中应用 non_padding_mask 过滤 padding, 提升监控准确性。
- 关注依赖兼容性: NumPy 2.0 迁移后, 监控 NPU 和 ROCm 后端是否出现兼容问题, 必要时设置环境降级。