

PR #6103 完整报告

verl-project/verl

[ci] fix: remove spmd test

合并时间: 2026-04-22 10:32

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6103>

执行摘要

- 一句话: 删除 FSDP 与 vLLM 集成的 SPMD 测试脚本, 清理测试目录。
- 推荐动作: 此 PR 变更简单, 无需精读。值得关注的点是团队在清理测试文件时的决策, 反映了对仓库结构的维护。建议工程师了解该脚本的删除原因, 以防未来需要类似集成示例。

功能与动机

PR 标题和提交信息直接说明“remove spmd test”, 表明目的是清理测试目录。PR body 中未提供具体原因, 但从上下文推测, 该测试脚本可能已过时、冗余或不再维护, 因此被移除以简化仓库。

实现拆解

1. 删除测试脚本: 移除了文件 `tests/workers/rollout/rollout_vllm/run_fsdp_vllm.py`, 该文件包含一个完整的测试示例, 演示了 FSDP 与 vLLM 的集成。
2. 清理测试配套: 删除后, 相关的测试覆盖、导入关系和控制流调整也随之移除, 但未涉及其他测试或配置文件的改动。
3. 无新增或修改: 此 PR 仅执行删除操作, 未添加新文件或修改现有代码, 因此无需测试、配置或部署配套调整。

关键文件:

- `tests/workers/rollout/rollout_vllm/run_fsdp_vllm.py` (模块 测试脚本; 类别 test; 类型 deletion; 符号 `_pre_process_inputs`, `main`): 这是唯一被删除的文件, 包含 FSDP 与 vLLM 集成的测试示例, 移除后影响测试覆盖。

关键符号: `_pre_process_inputs`, `main`

关键源码片段

`tests/workers/rollout/rollout_vllm/run_fsdp_vllm.py`

这是唯一被删除的文件, 包含 FSDP 与 vLLM 集成的测试示例, 移除后影响测试覆盖。

此文件已被完整删除, 原内容如下 (整理后):

```
import os
```

```
import torch
```

```
from torch.distributed.fsdp import FullyShardedDataParallel as FSDP
```

```

from transformers import AutoTokenizer, AutoModelForCausalLM
from vllm import SamplingParams
from verl.third_party.vllm import LLM
from verl.utils.distributed import initialize_global_process_group

def _pre_process_inputs(pad_token_id, prompt_token_ids: torch.Tensor) -> list[int]:
    """移除左填充token，以便将提示输入vLLM。"""
    non_pad_index = torch.nonzero(prompt_token_ids != pad_token_id, as_tuple=False)[0][0]
    return prompt_token_ids[non_pad_index:].tolist()

def main():
    # 初始化分布式环境
    local_rank, rank, world_size = initialize_global_process_group()

    # 加载模型和 tokenizer
    tokenizer = AutoTokenizer.from_pretrained(local_model_path, trust_remote_code=True)
    actor_model = AutoModelForCausalLM.from_pretrained(local_model_path, trust_remote_code=True)
    actor_model.to(torch.bfloat16)

    # 准备输入数据
    prompts = ["The president of the United States is", "The capital of France is", "The future of AI is"]
    input_ids = tokenizer(prompts, return_tensors="pt", padding=True)["input_ids"]

    # 使用 FSDP 包装模型
    fsdp_model = FSDP(actor_model, sharding_strategy=ShardingStrategy.FULL_SHARD, mixed_precision=MixedPrecision(param_dtype=torch.bfloat16))

    # 同步状态字典
    FSDP.set_state_dict_type(fsdp_model, state_dict_type=StateDictType.SHARDED_STATE_DICT)
    state_dict = fsdp_model.state_dict()

    # 配置 vLLM 采样参数
    sampling_params = SamplingParams(temperature=0, max_tokens=32)

    # 初始化 vLLM 引擎并生成响应
    llm = LLM(model=None, tokenizer=tokenizer, model_hf_config=actor_model_config, tensor_parallel_size=4)
    outputs = llm.generate(prompts, sampling_params)

    print(f"Generated responses: {outputs}")

```

评论区精华

Review 中无实质性讨论。gemini-code-assist[bot] 的评论仅指出“此 PR 移除了一个示例脚本”，并说明“没有反馈可提供”。wuxibin89 直接批准，未提出任何问题或疑虑。这表明变更被团队接受，无需深入讨论。

- 暂无高价值评论线程

风险与影响

- 风险：风险较低，主要涉及测试覆盖减少：
 - 回归风险：删除的脚本是一个演示性测试，不涉及核心业务逻辑，因此不会影响生产代码。
 - 兼容性风险：无，因为未修改任何 API 或依赖。
 - 安全风险：无，仅删除文件。
 - 性能风险：无，不影响运行时性能。潜在风险是如果该脚本被其他测试或文档引用，可能导致链接失效，但 PR 上下文中未发现此类依赖。
- 影响：影响范围有限：
 - 对用户：无直接影响，因为这是内部测试脚本，不暴露给终端用户。
 - 对系统：无功能变更，系统行为保持不变。
 - 对团队：简化了测试目录，减少了维护负担，但可能移除了一个有用的集成示例，需确保有替代文档或测试。
 - 风险标记：测试覆盖减少

关联脉络

- PR #6102 [ci] chore: update npu docker build pipeline: 同属 CI 相关清理或调整，涉及测试和配置优化。
- PR #6043 [ci] chore: change some npu ci test yml machine: 均涉及 CI 测试配置的修改，反映团队在持续优化测试基础设施。