

PR #3181 完整报告

radixark/miles

docs: DeepSeek-V4.1 Flash cookbook

合并时间: 2026-09-10 14:12

原文链接: <http://prhub.com.cn/radixark/miles/pull/3181>

执行摘要

- 一句话: 新增 DeepSeek-V4.1 Flash 训练配方文档并接入导航
- 推荐动作: 值得精读, 尤其是 §3.5 的必选 flag 解释与 §3.6 的预期指标表——它们把 colocated 引擎休眠释放 engram 表、全参数 Adam 状态必须落盘这些隐性坑写成了显式约束, 是理解 V4.1 训练栈成本结构的最短路径。建议跟进三件事: ① 明确 dsv41_cut_cast.py 的仓库路径或提供等价工具 (回应 issue 评论); ② 确认 docs.json 中 developer/ci/* → ci/* 的回退是否有意, 必要时拆成独立 PR; ③ 为文档中的脚本与参数增加轻量可用性检查, 避免 cookbook 成为新的答疑入口。

功能与动机

PR body 为空, 动机可从内容与关联讨论推断: DeepSeek-V4.1 Flash 是一个已在 miles 上跑通 (Megatron 插件位于 miles_plugins/models/deepseek_v41/, 启动器为 scripts/run_deepseek_v41.py) 但缺少公开复现路径的新模型, 需要一个页面把镜像、checkpoint 转换、16 卡编排、必选 flag 与预期指标一次性讲清。文档中“Three flags are not optional in this configuration”一节说明作者踩过真实坑 (engram host table 缺失会让 train/train_rollout_logprob_abs_diff 从 ~0.02 跳到 ~10), 写文档的目的就是把这些隐性知识固化。issue 评论 DanielSun94 的提问“mentions using dsv41_cut_cast.py to convert the model weights, but I couldn't find the script”则反映出文档已经进入实际使用环节。

实现拆解

1. 新增配方页 (docs/models/deepseek/deepseek-v4-1-flash.md, +225行): 主体分三块。
 - ① 模型结构: model_type: deepseek_v41 (兼容旧名 deepseek_v4.1)、40 层解码器、384 个路由专家 top-6 加 1 个共享专家、dense 权重 fp8 (32 宽 ue8m0 block scale) + 路由专家 fp4; 每层用 512 宽 KV latent 同时充当 key/value, 64 个 query head 无独立 value 投影; 压缩 latent 在 2-19 层做 2 合 1、20 层起 1 比 1, 另有每层重算、不随上下文增长的 128 位置滑窗; 残差流为四路并行副本, 由 per-token 双随机矩阵混合。
 - ② 稀疏检索: 强调 KV source layer (4 个) 与 index source layer (8 个) 是两份不同列表, 检索层选 top-512 候选, 20 层起有粗粒度候选块预筛。
 - ③ Engram: 两层上的加性 n-gram 哈希记忆, token id 先归一化再哈希, 其两张 fp8 表是 checkpoint 中最大的单块权重。训练侧落点写明 Megatron 插件在 miles_plugins/models/deepseek_v41/、模型定义 scripts/models/deepseek-v4.1.py、启动器 scripts/run_deepseek_v41.py, 插件在训练前向复现 fp4/fp8 量化点, 并把跨层状态 (混合系数、source latent、top-k、候选掩码) 装

进层间 hidden tensor 以兼容流水并行与全量重算。

2. Quick Start 流程 (§3.1-§3.5) : 拉专用镜像 radixark/miles:deepseek-v41 (pin 版本记在 /sgl-workspace/sglang/.sglang_rev、/root/miles/.miles_rev 与镜像标签) → 准备 BF16 HF checkpoint (fp8 dense 与 fp4 专家反量化到 bf16, 两张 engram 表保留 fp8 bits + e8m0 scale 约 197 GB, 丢弃 MTP/DSpark/vision 张量, 产出约 1.2 TB 且必须落在每个训练节点的本地 NVMe) → prepare-data --task dapo_aime → 手动 ray start 组 16 卡集群 → head 节点提交 scripts/run_deepseek_v41.py train。
3. 必选 flag 与预期指标 (§3.5、§3.6) : 用表格单独解释三个不可省参数
——--sglang-engram-host-table (colocated 引擎休眠会释放全部 GPU 状态, engram 表是常量且不在权重更新链路里, 不共享 host 表就会读回空表)、--disk-offload
--offload-disk-dir (全参数 Adam 状态在 16 卡上放不下, 逐步流式写 NVMe)、
--load-from-hf (bf16 checkpoint 由 mbridge 启动期映射, 省掉多机 torch_dist 转换)。
§3.6 给出可核对基线: step 0 的 train/train_rollout_logprob_abs_diff ~0.02 nats、
train/train_rollout_kl 0.0012-0.0017、rollout/raw_reward ~0.5 并在 80 步内升过 0.75、
每步约 11 分钟、训练峰值显存约 218 GB/277 GB。
4. 导航与索引联动: docs/docs.json 的 Models tab 在 DeepSeek 分组下新增
DeepSeek-V4.1 子组, 指向新页面; docs/models/deepseek/index.md 把“两代”文案改为“三代”并在 Variants 表新增 V4.1 Flash 行; docs/models/index.md 的 DeepSeek 行补上
V4.1 Flash 链接; 新增配图 docs/assets/images/dsv41-rl-training.png 供配方页引用。
5. 配套与缺口: 没有任何测试或 CI 校验覆盖这批文档; docs/docs.json 同时把 Developer Guide 下的 CI 页面从 developer/ci/* 改回 ci/* 并删除 7 条 /ci/* → /developer/ci/* 的 permanent redirect, 与本 PR 主题无关, 属于需要确认动机的附带改动。§3.2 提到的 dsv41_cut_cast.py 未随 PR 入库, 文档只写“kept with the bring-up scripts”。

关键文件:

- docs/models/deepseek/deepseek-v4-1-flash.md (模块 模型文档; 类别 docs; 类型 documentation) : 本 PR 的主体, 新增 DeepSeek-V4.1 Flash 的端到端训练配方, 覆盖模型结构、checkpoint 转换、16 卡启动命令、三个必选 flag 与预期指标基线, 是用户复现该模型的唯一入口。
- docs/docs.json (模块 文档导航; 类别 config; 类型 configuration) : 文档站导航契约: 新增 DeepSeek-V4.1 分组使新配方页可达, 但同时把 Developer Guide 下 CI 页面路径从 developer/ci/改回 ci/ 并删除对应的 permanent redirect, 是本次唯一的非预期 / 需确认改动。
- docs/models/deepseek/index.md (模块 模型索引; 类别 docs; 类型 documentation) : DeepSeek 家族总览页, 把代数描述从“两代”改为“三代”并在 Variants 表新增 V4.1 Flash 行, 是新页面的主要发现入口。
- docs/models/index.md (模块 模型索引; 类别 docs; 类型 documentation) : 模型总览页, DeepSeek 行补上 DeepSeek-V4.1 Flash 链接, 保证新页面可从顶层索引进入。
- docs/assets/images/dsv41-rl-training.png (模块 文档配图; 类别 other; 类型 asset) : 为配方页新增的配图 (RL 训练流程示意), 属文档资产, 无代码语义。

关键符号: 未识别

关键源码片段

docs/models/deepseek/deepseek-v4-1-flash.md

本 PR 的主体，新增 DeepSeek-V4.1 Flash 的端到端训练配方，覆盖模型结构、checkpoint 转换、16 卡启动命令、三个必选 flag 与预期指标基线，是用户复现该模型的唯一入口。

```
# 在 head 节点执行。验证配置：4 节点 x 4 张 GB300（16 张 GPU），colocated
rollout，训练与推理均为 BF16
# MILES_SCRIPT_EXTERNAL_RAY=1：复用 S3.4 手动拉起的 Ray 集群，而不是由脚本自建
#
# 文档单独列表强调的“三个不可省 flag”：
# --sglang-engram-host-table colocated 引擎休眠会释放全部 GPU 状态，engram 表是常量、
# 不参与权重更新；没有共享 host 表就会被读回空表，
# 使 train/train_rollout_logprob_abs_diff 从 ~0.02 跳到 ~10
# --disk-offload/--offload-disk-dir 全参数 Adam 状态在 16 张 GPU 上放不下，
# 逐步流式写到节点本地 NVMe
# --load-from-hf BF16 HF checkpoint 由 mbridge 启动期映射，
# 省掉多机 torch_dist 转换
#
# 环境耦合点：/scratch 需在每个节点同路径存在，eth0 需替换为集群实际网卡
MILES_SCRIPT_EXTERNAL_RAY=1 python scripts/run_deepseek_v41.py train \
  --model-name DeepSeek-V4.1 \
  --hf-checkpoint /scratch/models/DeepSeek-V4.1-bf16 \
  --model-dir /scratch/models --data-dir /scratch/datasets \
  --num-nodes 4 --num-gpus-per-node 4 --hardware GB300 \
  --task dap0_aime --mode debug_minimal \
  --pp-size 2 --recompute full \
  --load-from-hf \
  --disk-offload --offload-disk-dir /scratch/offload \
  --colocate-memory-peak-device cpu \
  --sglang-engram-host-table --sglang-cuda-graph --sglang-radix-cache \
  --no-check-weight-update \
  --extra-args "--stream-optimizer-state-moment-dtype bf16"
```

docs/docs.json

文档站导航契约：新增 DeepSeek-V4.1 分组使新配方页可达，但同时把 Developer Guide 下 CI 页面路径从 developer/ci/改回 ci/ 并删除对应的 permanent redirect，是本次唯一的非预期 / 需确认改动。

```
// Models tab: 在 DeepSeek 分组下新增 DeepSeek-V4.1 子组，顺序排在 DeepSeek-V4 之前
{
  "group": "DeepSeek",
  "pages": [
    "models/deepseek/index",
    {
      "group": "DeepSeek-V4.1",
      "pages": [
        "models/deepseek/deepseek-v4-1-flash" // 必须与新增 md
        文件路径严格一致，否则构建期报错
      ]
    }
  ]
}
```

```

    ],
    "expanded": false
  },
  {
    "group": "DeepSeek-V4",
    "pages": [
      "models/deepseek/deepseek-v4-flash",
      "models/deepseek/deepseek-v4-pro"
    ],
    "expanded": false
  }
]
}

```

// Developer Guide tab: CI 页面路由由 developer/ci/* 改回 ci/* (base 侧为 developer/ci/*) ,
 // 同时 redirects 段里 7 条 "/ci/* -> /developer/ci/*" 的 permanent 跳转被整段删除。
 // 这两处与 V4.1 cookbook 无关, 属于夹带的路径回退, 需要向作者确认是否为误操作。

评论区精华

本 PR 的 review 讨论极轻: 只有 claude[bot] 的手动 review 配置提示和 yueming-yuan 的一次无评论 APPROVED, 没有 diff 级别意见, 也没有人指出 docs.json 中 CI 路径被改回 ci/* 这件事。真正有信息量的是 issue 评论: DanielSun94 指出“the cookbook for deepseek-v4-1-flash in this commit mentions using dsv41_cut_cast.py to convert the model weights, but I couldn’t find the script. Could you please let me know where it is located?”, 作者未在此 PR 中给出仓库路径或替代获取方式, 属于明确未解决的文档闭环缺口。

- 文档引用的 dsv41_cut_cast.py 在仓库中不存在 (documentation): 作者/维护者尚未回复或补充链接; 文档给出的转换路径目前对仓库外用户不可执行。
- docs.json 中 CI 文档路径被改回 ci/* 且未在 review 中讨论 (design): 缺少明确结论; 需要与 CI 文档维护者确认是有意回退还是分支陈旧导致的夹带改动。

风险与影响

- 风险:
 1. 文档引用脚本缺失 (已在真实使用中暴露): §3.2 指定用 dsv41_cut_cast.py --src <fp8 ckpt> --dst <bf16 ckpt> --config <flattened config.json> --layers 0,...,39 做 fp8/fp4/engram 混合转换, 但仓库内只有覆盖 fp8 dense 的 tools/fp8_cast_bf16.py, 用户按文档操作会在第一步卡住; issue 评论已直接印证。
 2. docs.json 导航回退风险: Developer Guide 下 7 个 CI 页面从 developer/ci/* 改回 ci/*, 同时删除 7 条 permanent redirect。若这是误回退, 已有外链 (/ci/*) 与新路径的映射会丢失并导致 404; 若是刻意回退, 则本 PR 混入了与前一个路径迁移 PR 相矛盾的改动, 建议与 CI 文档负责人确认。
 3. 文档易腐且无自动校验: 启动命令里的 --no-check-weight-update、--sglang-engram-host-table、--stream-optimizer-state-moment-dtype bf16 与 mode debug_minimal 都会随代码演进失效, 文档没有任何链接 / 参数可用性检查或文档构建

测试兜底。

4. 环境强耦合与高门槛：/scratch、eth0、4x4 GB300、镜像标签 deepseek-v41 全部硬编码；1.2 TB checkpoint 加约 1 TB 优化器状态的分片要求，读者误读成本较高。 - 影响：对用户：DeepSeek-V4.1 Flash 首次有了单页可复现的 BF16 RL 配方，§3.6 的预期指标 (abs_diff ~0.02 nats、raw_reward 0.5→0.75、每步 11 分钟) 让使用者能自行判断是否复现成功，价值明显。对系统：零运行时代码改动，无回归风险，但会提高该模型与 radixark/miles:deepseek-v41 镜像的使用量，把文档与脚本分发不一致的问题一并放大。对团队：短期增加支持负担（脚本缺失的答疑），长期需要一套文档 / 命令可用性校验机制；同时暴露了文档 PR 因 review 轻量化而漏检导航回退类问题。 - 风险标记：文档引用脚本缺失，导航路径疑似回退，无测试覆盖，命令与配置易腐

关联脉络

- PR #2038 [feat] Enable THD packed-sequence training for DeepSeek-V4-Flash, with context parallelism: 同属 DeepSeek-V4 模型线的训练栈改造，本 PR 描述的跨层状态塞入层间 hidden tensor 以兼容流水并行 / 全量重算，正是该系列工程约束的延续，二者需对照阅读才能理解 V4.1 与 V4 的差异。
- PR #3162 ci: pin the DSv4 bshd/thd parity test to the miles impl: DeepSeek-V4 家族已有 bshd/thd 精度对等测试作为 CI 护栏，而本次新增的 V4.1 Flash 目前只有文档、没有 e2e 或精度测试，形成同一模型线内的覆盖落差。
- PR #2855 Restrict NVFP4 RL to routed experts: 同样围绕“哪些权重该量化、哪些必须保持高精度”的取舍；本 PR 文档中 fp8 dense + fp4 路由专家 + engram 保留 fp8 bits 的混合精度结构，与该 PR 收紧量化范围到路由专家的结论属同一设计脉络。