

PR #32120 完整报告

sgl-project/sglang

[AMD][DI][CI] 8/N Add GLM-5.2 MXFP4 1P1D DI/CI recipes (base + MTP + DP8/EP8)

合并时间: 2026-08-07 13:56

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/32120>

执行摘要

- 一句话: MI355X 新增 GLM-5.2 MXFP4 四类 1P1D 夜测 recipe
- 推荐动作: 值得快速浏览以了解 AMD MI355X CI recipe 的结构与 launcher 路径约定, 但无需精读, 因为全部为配置变更。更值得关注的是被暴露的 DSA + NextN + DP-attention 崩溃 bug, 建议跟踪其后续修复 PR, 并留意 model_path 从 5.1 切换到 5.2 checkpoint 的跟进 PR。合并决策本身合理: 3/4 腿验证通过且改动不触碰现有配置。

功能与动机

让 AMD MI355X 的 nightly 以与 DeepSeek-V4、Kimi-K2.6 相同的方式覆盖 GLM MoE 模型线。PR 描述原文: Add AMD MI355X DI/CI nightly recipes for GLM-5.2 (MXFP4) 2-node 1P1D disaggregation, so the nightly exercises the GLM MoE line the same way it already covers DeepSeek-V4 and Kimi-K2.6。GLM-5.2 使用 GlmMoeDsaForCausalLM (DSA 稀疏注意力, sglang 自动选择后端) 且内置 NextN MTP 头 (num_nextn_predict_layers=1), 无需外部草稿 checkpoint, 适合纳入现有 1P1D basic tier 测试矩阵。

实现拆解

1. 新增 4 个 recipe 文件 (scripts/ci/slurm/recipes/mi355x-fp4/glm52/1k1k/ 下):
1p1d.yaml (TP8 1P1D 无 MTP)、1p1d-mtp.yaml (TP8 + 内置 NextN MTP)、
1p1d-dp8ep8.yaml (DP-attention 8 + 窄 EP8)、1p1d-dp8ep8-mtp.yaml (DP8 + EP8 + NextN MTP)。每个文件声明 resources (prefill/decode 各 1 worker)、
backend.sglang_config (TP/EP/DP) 与 bench 段 (并发扫描 + GSM8K accuracy gate, threshold 0.91)。
2. model 块与通用 launcher 路径: 4 个 recipe 均携带 model 块 (SGLANG_USE_AITER=1 环境变量 + server_args), 其存在与否决定 launch_mi355x.sh 走通用命令构建路径而非硬编码 DeepSeek-V4 分支。后者会无条件追加 --attention-backend, 对 GLM 会报 expected one argument 错误。GLM 有 n_shared_experts=1, 因此显式传入 --disable-shared-experts-fusion; attention_backend 故意不设置, 让服务端对 GlmMoeDsaForCausalLM 自动选择 DSA 后端。
3. nightly-configs.yaml 注册 4 个执行块: 追加 glm52-fp4-mi355x[-mtp1-dp8ep81-dp8ep8-mtp]-sglang, 统一配置 model、model-prefix、runner=mi355x、precision=fp4、

framework=sglang、multinode=true、disagg=true、seq-len 1k/1k 与搜索结果映射到上述 recipe 文件。model_path 指向 /it-share 上镜像的 amd/GLM-5.1-MXFP4 checkpoint（见风险）。

4. 镜像与 MTP 修复依赖：recipes 引用 lmsysorg/sglang-rocm:v0.5.15.post1-rocm720-mi35x-20260722，该镜像包含 #30265（GlmMoeDsaForCausalLMNextN 类 + GLM 专属 Quark excludes / MTP layer-index remap），是 GLM NextN MTP 精度的前置修复。提交历史显示作者先修 launcher 启动问题、再升级镜像并完成 4 腿验证。
5. 验证配套：在 MI355X 2 节点 1P1D 上以完整 GSM8K 1319 题 8-shot 验证，base 0.930、mtp 0.933、dp8ep8 0.945 均超过 gate 0.91；dp8ep8-mtp 在 DSA topk indexer 崩溃，作为已知源码 bug 单独跟踪。

关键文件：

- scripts/ci/slurm/recipes/mi355x-fp4/glm52/1k1k/1p1d-mtp.yaml（模块 CI 配置；类别 infra；类型 infrastructure）：TP8 + 内置 NextN MTP 的 recipe，展示 model 块、mtp 段与 accuracy gate 的完整配置，是理解通用 launcher 路径的代表文件。
- scripts/ci/slurm/recipes/mi355x-fp4/glm52/1k1k/1p1d.yaml（模块 CI 配置；类别 infra；类型 infrastructure）：基础 TP8 1P1D recipe，是所有变体的模板，也是验证通过（GSM8K 0.930）的第一个腿。
- scripts/ci/slurm/recipes/mi355x-fp4/glm52/1k1k/1p1d-dp8ep8.yaml（模块 CI 配置；类别 infra；类型 infrastructure）：DP-attention 8 + 窄 EP8 拓扑，验证通过（0.945），与 DeepSeek-V4 dp8ep8 腿同构。
- scripts/ci/slurm/recipes/mi355x-fp4/glm52/1k1k/1p1d-dp8ep8-mtp.yaml（模块 CI 配置；类别 infra；类型 infrastructure）：DP8 + EP8 + NextN MTP 组合腿，当前因 DSA topk 崩溃无法通过，是暴露源码 bug 的关键线索。
- scripts/ci/slurm/nightly-configs.yaml（模块 夜测注册；类别 infra；类型 infrastructure）：nightly 执行注册中心，追加 4 个 GLM-5.2 块，将 recipe 映射到模型、精度与搜索空间。

关键符号：未识别

评论区精华

该 PR 没有实质性的 review 评论，HaiShaw 直接批准（Approved，body 为空）。最有价值的技术讨论来自 PR 作者自述的 dp8ep8-mtp 崩溃分析，被定位为源码级缺陷而非 recipe 问题：

```
RuntimeError: Expected lengths.size(0) == B to be true, but got false in the DSA
sparse-attention top-k indexer during the NextN draft forward, only when
DP-attention and MTP are combined。调用栈为 deepseek_nextn.py:276 →
deepseek_v2.py:1860 → forward_mla.py:384 → dsa_indexer.py:1998 →
dsa_backend.py:315 → sgl_kernel/top_k.py:77 fast_topk_transform_fused。
```

结论：DP-attention 与 MTP 单独使用均正常（dp8ep8 = 0.945，MTP = 0.933），组合时 per-batch seqLens 未按 DP-padded NextN draft-extend batch 构建，属 DSA + NextN + DP-attention 的源码 bug，超出 recipe PR 范围，单独跟踪。作者明确要求不得将该腿视为通过。

- dp8ep8-mtp 腿在 DSA topk indexer 崩溃 (correctness): 判定为 DSA + NextN + DP-attention 组合的源码级 bug, 非 recipe 配置问题, 超出本 PR 范围, 单独跟踪; 该腿保留在 recipes 中但不得视为通过。

风险与影响

- 风险:
 1. dp8ep8-mtp 腿已知崩溃但已合入: 该 recipe 与 nightly 块保留在仓库中且标记为不通过, 若后续 nightly 自动执行该块将持续报红, 可能被当作已知失败忽略, 从而掩盖同路径上的新回归。
 2. checkpoint 命名与内容不一致: recipe/block 命名跟踪 GLM-5.2 模型线, 但 model_path 指向磁盘上唯一的 amd/GLM-5.1-MXFP4 镜像; 若 5.2 checkpoint 的 MoE 布局或 DSA 行为与 5.1 不同, 验证结果会产生误导, 作者已声明 5.2 checkpoint 落地后需重定向。
 3. 依赖特定 nightly 镜像: 镜像 20260722 包含 #30265 修复, 是 MTP 精度前提; 镜像更新或回退会直接影响这 4 个腿的结果, 且没有自动化断言校验镜像内容。
 4. 纯配置无源码断言: 没有针对 recipe 结构或 launcher 路径的测试, 未来 launch_mi355x.sh 的行为变更 (如通用路径分支调整) 可能悄悄破坏这些 recipe。
 5. 硬件环境假设: ib_devices rdma0-3、端口号、MORI 网络拓扑为特定集群假设, 环境变化时 recipe 需要人工同步。- 影响: 影响范围限于 CI/nightly 基础设施: 为 AMD MI355X 分解式推理新增 GLM MoE 模型线覆盖, 使 nightly 与 DeepSeek-V4、Kimi-K2.6 矩阵对齐, 可更早捕获 GLM-5.2 在 ROCm 平台的精度与性能回归。对运行时、用户 API 无影响。对团队而言, 新增 4 个执行块会占用 nightly GPU 资源, 其中 1 个腿当前必然失败, 需要运维侧按已知失败处理。整体影响程度为中低, 价值在于覆盖缺口与后续源码 bug 的跟踪线索。- 风险标记: dp8ep8-mtp 已知崩溃未修复, checkpoint 命名与内容不一致, 依赖特定 nightly 镜像版本, 纯配置变更无自动化断言

关联脉络

- PR #30265 [AMD] Fix GLM-5.2 MTP Quark excludes: PR body 明确引用: 镜像 20260722 包含该修复 (GlmMoeDsaForCausalLMNextN 类 + Quark excludes / MTP layer-index remap), 是 GLM NextN MTP 精度与可用性的前置依赖。
- PR #33935 Clean GLM-5.2 NVFP4 cookbook: 同为 GLM-5.2 模型线的配套文档变更, 说明 GLM-5.2 在仓库中的功能落地脉络。
- PR #32466 [AMD] Enable gfx1250 sgl-kernel builds: 同为 AMD 平台基础设施变更, 与 MI355X 系列 CI 覆盖同属 AMD 夜间测试建设方向。