

高梓源

(+86) 1980-119-0365 · gaosion2001@gmail.com · GitHub @ETOgaosion



教育背景

中国科学院大学, 计算机科学与技术, 本科 2019.9 - 2023.6

- GPA: 3.8/4.0, 排名: 18/103(前 17%), 中国科学院大学三好学生 (2 次), 中国科学院大学三等奖学金 (3 次)
- CET-4: 587, CET-6: 570

中国科学院计算技术研究所, 计算机科学与技术, 硕士研究生 2023.9 - 2026.6

- 联想企业冠名奖学金

技术能力

- 大模型训练推理系统: Megatron-LM, Deepspeed, 了解 vLLM, SGLang, 有大模型后训练系统 (verl) 开发经验
- 大模型训练优化技术: Hybrid Parallelism, Flash Attention, ZeRO-1~3, Torch FSDP + CPU Offload
- 大模型推理优化技术: KVCache, 了解 Paged Attention, Prefill-Decode 分离, 量化技术
- 编程语言: C/C++(STL, C++11, openmp), Python, Golang, SQL, 了解 CUDA/Triton

实习经历

字节跳动 (Seed-Infra-训练), 大模型后训练系统实习生, [volcengine/verl](#) (12K stars) 2025.2-至今

- verl 开源核心开发者 (近 100 个 PR 被采用), 支持 verl 使用最新版 Megatron 进行后训练, 并支持 Qwen3-235B 和 DeepSeek-V3 671B 等大 MoE 模型的后训练及性能优化
- 功能支持: Megatron 高版本 GPTModel 接入与 Sequence Packing 功能对齐, 支持模型隔离的 5D 并行与 MoE Parallel Fold 技术, RL 侧算法支持且与 FSDP 收敛性对齐 (MoE 模型配置细节), Megatron checkpoint 支持及其与 huggingface model 的分布式流式转换 (后训练新需求), 实现后训练多进程场景下基于 Ray 的多种 Profiler
- 显存优化 (DeepSeek-671B Infra): 全流程节省 CPU-GPU Memory
 - 初始化: 训练推理引擎初始化, 第一次模型加载
 - Rollout Generation: Megatron 与推理引擎权重高效同步, 推理引擎卸载
 - 训练: Megatron 完全 CPU Offload, Fused Linear/LogProb/Cross-Entropy Kernel 接入 Megatron 削减显存峰值, CPU Adam 优化器
- 性能优化: 目前以训练侧 MFU 为目标, H20 上 MFU(Dense Model)>0.8, MFU(MoE Model)>0.4, H100 上 Qwen3-MoE(235B) 达到 0.22 的预训练水平
 - Kernel Fusion: Packed Sequence 场景必需 rope fusion, TP backward 时 gradient accumulation fusion, moe permute/unpermute 操作等
 - 计算通信重叠: 挑战是通信会影响正常计算且不成比例, NCCL 通信操作需要双方同步等待; 具体来说在 TP/SP 进行计算/通信 Overlap, DDP 时将 gradient 计算与 all-reduce Overlap, param forward 计算与 gather 操作及 Optimizer Step 进行 Overlap
 - MoE 优化: grouped GEMM 与 deep-ep, shared experts 计算与 routed experts dispatch 通信重叠
 - 后训练 Challenge: 序列长度更加不均, 需要 dynamic sequence balancing (Karmarkar_karp 算法)
 - 高效自动并行配置: (3*2)+2D 并行 (MoE Parallel Folding 技术)+ 重计算粒度, 并且在后训练中不同角色显存空间不同、随着训练过程序列长度不断变化, 静态单配置低效
- 异步 RLHF 及长尾: 课程学习与生成长度预测器, 减少训练等待时间, 使训练负载更加均衡。Semi-Sync RLHF

爱诗科技, 大模型训练系统实习生 2024.7-2024.9

- 独立负责 Unified Sequence Parallel(Deepspeed Ulysses + Ring Attention) 在 Megatron-LM 中的实现, 支持多模态长序列训练 (128K), 并完成训练精度对比
- 迁移 Deepspeed 中 T2I, T2V, DiT 模型及算子至 Megatron, 支持「3D」并行 (TP/CP/DP)

华为云计算技术有限公司, 科研实习生 2024.9-2024.10

- 负责探究在昇腾异构平台上的自动并行配置方案, 完成在 Pipeline 维度的算子级灵活切分, 并在不同异构设备组上支持不同的 TP/CP/DP/重计算配置, 同时支持数据并行组间 batch size 的动态调整
- 性能相比同构 Megatron 方案提升 24%, 相较于异构自动配置工作 Metis 能够处理更长文本序列

研究工作

大模型分布式弹性训练与容错系统, 论文在投 2023.6-2024.10

- 在大模型分布式训练中, 故障的处理和恢复过程会导致训练进度的丢失和训练效率的下降, 特别是使用相对廉价的 AWS Spot Instance 集群与训推混部集群
- Megatron-LM 中的规整 3D 并行在出现节点故障时, 原有并行方案失效, 缺乏弹性训练能力。
- 本项目实现了自适应流水线, 允许自由的流水线 Stage 切分方案, 节点故障后剩余节点能够以高效的配置续训
- 将模型的参数、优化器参数和每个 micro batch 的梯度做冗余备份到其他节点, 该过程与计算 Overlap, 并针对不同网络和算力条件设计了多种备份方案, 在保证计算性能的前提下实现 micro batch 级的故障恢复粒度
- 实现更接近真实的模拟器, 能够模拟更长时间训练吞吐量变化, 基于 AWS 竞价实例数量变化的 Trace 和可控节点 Fail 频率进行仿真, 证明本方案性能优于 Bamboo, Varuna, Oobleck 等工作