GD2FS产品手册 v20251104

GD2FS简介

GD2FS（GPU Direct Distributed File System）是面向下一代AI基础设施构建的高性能分布式文件系统。其核心设计深度融合GPU直接访问与高速网络能力，为大语言模型训练、推理等高并发、低延迟场景提供极致的存储性能与显著的整体成本优势。

GD2FS具备高度灵活的部署能力，既支持基于RDMA的高速网络架构，实现超低延迟数据传输，也兼容TCP经典网络环境，保障广泛部署的便利性；既可在专用存储服务器上独立部署，也支持在AI计算节点上实现存算混合部署，为用户提供多样化的架构选择与优化路径。

核心特性

- GPU显存与远程存储直接数据交换：基于RDMA实现GPU显存与GD2FS之间的直接数据传输，数据全程零拷贝，逼近理论最优延迟，大幅降低CPU开销，1GB数据延迟可低至25毫秒。
- 高性能TCP多路并发：支持多网卡绑定与多队列并发I/O，1GB数据传输延迟可控制在70毫秒以内。
- 灵活的数据持久化策略：支持单副本存储，契合KV Cache可重复计算的特性，兼顾低延迟与低成本；同时提供多副本机制，确保模型文件与训练检查点数据高可靠存储。
- 智能缓存管理：支持1至N副本缓存策略，有效应对推理服务批量启动时可能出现的“启动风暴”。
- 高效的缓存调度机制：采用LRU淘汰策略，结合具备分布式语义的预读取（Readahead）与缓存释放（Drop Cache）能力，实现全局缓存的高效调度与利用率提升。

典型应用场景

- AI推理加速：实现KV Cache的高性能远程化读写，支持多轮对话场景下Cache完全命中，助力无状态推理架构，显著提升GPU利用率。
- 模型极速加载：大幅缩短模型加载时间，例如DeepSeek R1模型可在20秒内完成本地加载，减少GPU空闲等待，提升研发与推理效率。
- 训练任务高效执行：实现训练检查点（Checkpoint）的快速保存与恢复，有效压缩GPU运行中的“气泡时间”，提升训练任务整体吞吐。
- 训推一体架构支持：覆盖AI训练与推理全流程，为端到端AI业务链路提供统一的存储加速与资源优化。