

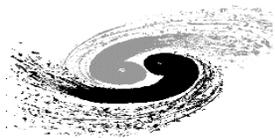


中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

四月份考核报告

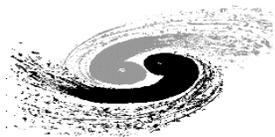
张 航 畅
导师：顾旻皓
计算机应用技术

2024. 04



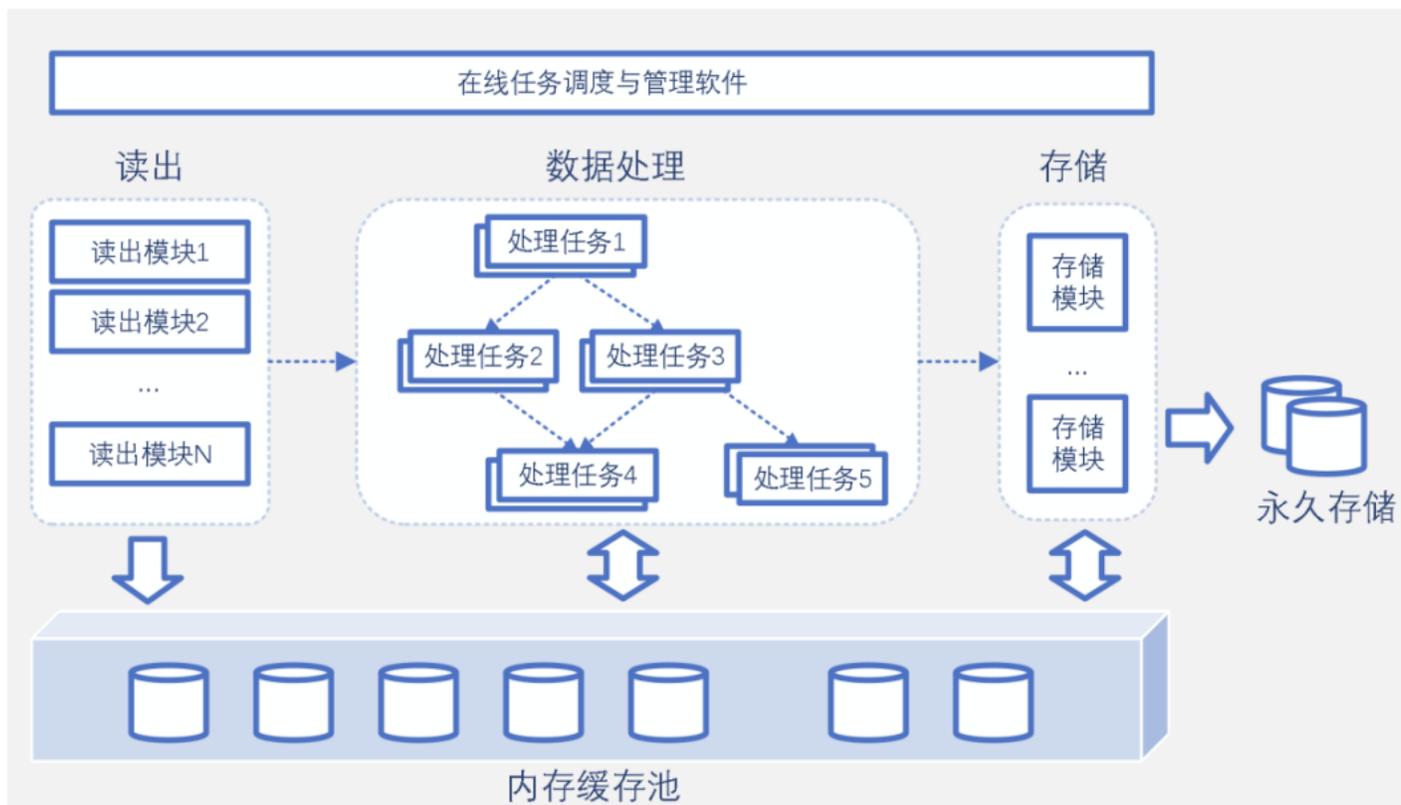
➤ 基于分布式内存缓存的在线计算方案

- ◆ 基于MINIO的分布式内存缓存池上线
- ◆ 开发缓存池读写接口
- ◆ 自研内存缓存池调研



基于分布式内存缓存的在线计算方案

- 基于文件的在线数据处理软件开发方案
- 以文件形式保存raw data于**内存缓存**中
- **在线处理框架**OPF读取文件进行数据处理，支持多条计算独立运行



基于分布式内存缓存的在线计算方案

基于MINIO的分布式内存缓存池上线

MINIO

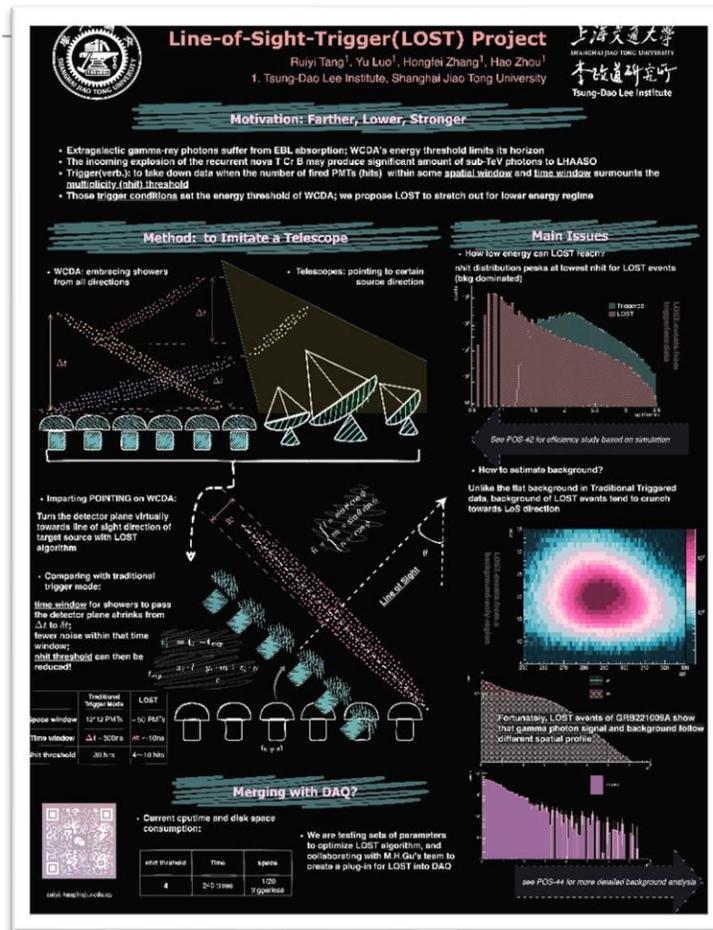
- MINIO 是个基于 Golang 编写的开源对象存储套件
- 可以存放任意类型的文件
- 容量和处理能力弹性扩展
- 在此之前，没有用于内存的先例

基于MINIO构建分布式内存缓存池：

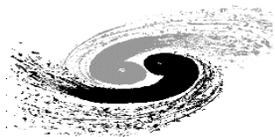
- 20台节点
- 7TB大小
- 已在KM2A Scalar 模式稳定运行数月

根据LOST触发模式要求，WCDA接入**内存缓存池**：

- ~1.7GB/s 数据率
- 34个RO节点
- 缓存30分钟数据



LOST软件触发模式



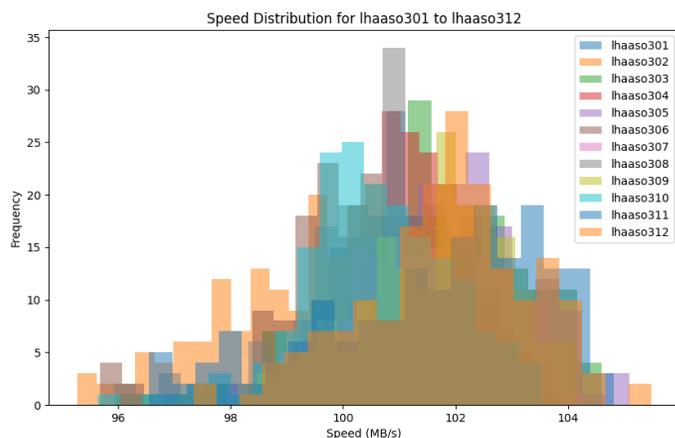
基于分布式内存缓存的在线计算方案

实现LOST触发模式

DAQ软件修改
RO代码

内存缓存池的
客户端性能优
化

为用户提供内
存缓存池交互
接口



使用多线程技术

- 将数据写入速度从40MB/s提高到了600MB/s (单节点)
- 多节点实现写入均值100MB/s, 满足WCDA要求

意外处理

- 忽略超过等待队列的数据
- 在MINIO连接失败时, 等待重连

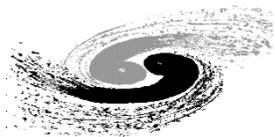
简单直接的接口

- `opf::write()`
- `opf::read()`

多种模式选择

提供用户手册和测试环境

满足LOST触发模式要求
为五月份观测做好准备

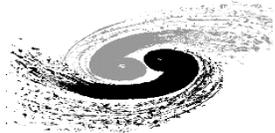


基于分布式内存缓存的在线计算方案

直接使用MINIO实际性能不及目标

➤ 调研与分析：

- ◆ 客户端效率低：HTTP接口
- ◆ 校验码耗时长
- ◆ 分布式事务处理耗时
- ◆ CPU利用率巨大

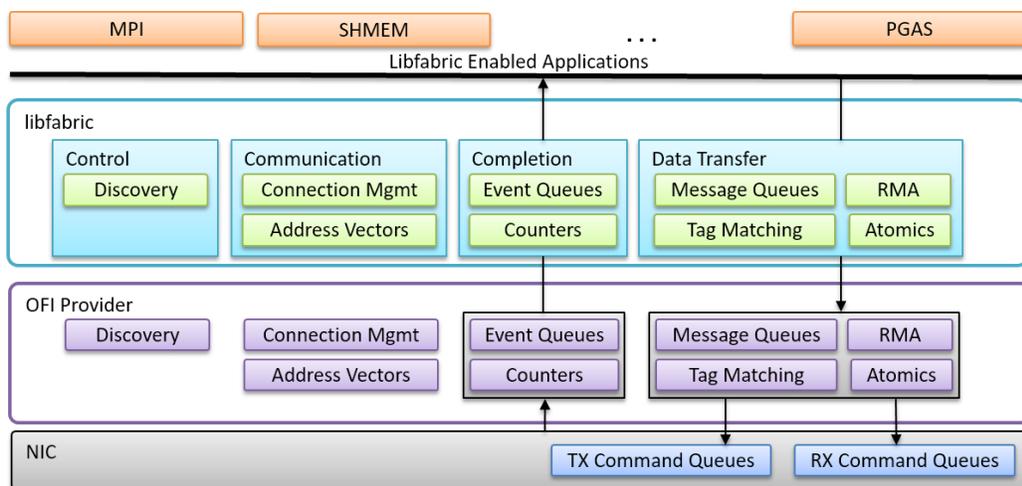


基于分布式内存缓存的在线计算方案

自研内存缓存池进展

网络层

◆ 基于RDMA的高性能通讯层 HYLIA: Hyper Link Access



基于 Intel Libfabric 库实现:

- 用于高性能并行和分布式应用程序的通信API
- 支持多种通信语义
- 支持几乎所有软件平台

已实现简单的RoCE demo
落地成产品还有一段距离

其他

➤ Member/合作组成员管理系统

◆ 开发超级管理员/管理员功能

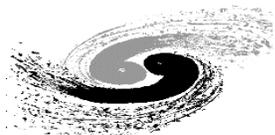
Actions	#	Email	First Name	Last Name	Inst.	Begin Time	End Time	Identity
	67	liwg@ihep.ac.cn	Weiguo	Li	IHEP	2014-07-28	-	User
Personal		mabikim@ihep.ac.cn	Medina	Ablikim	IHEP	2006-01-01	-	User
Group		baijz@ihep.ac.cn	Jingzhi	Bai	IHEP	2016-06-21	-	User
Inst		caixiao@ihep.ac.cn	Xiao	Cai	IHEP	2014-07-28	-	User
...	70	caixiao@ihep.ac.cn	Xiao	Cai	IHEP	2014-07-28	-	User
...	71	changjf@ihep.ac.cn	Jinfan	Chang	IHEP	2014-07-28	-	User
...	72	cheng@ihep.ac.cn	Gang	Chen	IHEP	2006-01-01	-	User
...	73	chenhs@ihep.ac.cn	Hesheng	Chen	IHEP	2006-01-01	-	User
...	74	chenjc@ihep.ac.cn	Jiangchuan	Chen	IHEP	2006-01-01	-	User
...	75	chenjin@ihep.ac.cn	Jin	Chen	IHEP	2006-01-01	-	User
...	77	chenyb@ihep.ac.cn	Yuanbo	Chen	IHEP	2006-01-01	-	User

➤ IEEE Real Time 2024 Poster

◆ Fault Detection and Diagnosis System for LHAASO



RT 2024 海报+MiniOral
TNS 论文初稿已完成



总结和下一步计划

➤ 基于分布式内存缓存的在线计算方案

- ◆ 基于**MINIO**的分布式内存缓存
- ◆ WCD A接入分布式内存缓存，满足新触发模式需求
- ◆ 开发读写接口，服务物理用户
- ◆ 基于RDMA的高性能网络通讯层的调研与开发

➤ 参与了IEEE Real Time 2024 会议

- ◆ Poster + MiniOral

➤ 下一计划

- ◆ 实现RDMA网络通讯层
- ◆ 实验运行故障诊断和管理系统 FADD
 - 实现落地
 - 1-2个月的工作量