



高海拔宇宙线观测站
Large High Altitude Air Shower Observatory



基于分布式内存缓存的在线计算方案 (LHAASO 在线数据处理软件升级)

张航畅 顾旻皓

中国科学院高能物理研究所

第二十一届全国核电子学与核探测技术学术年会

2023年8月10日

主要内容

- 背景
- 系统架构
- 应用
- 升级计划
- 总结



科学目标:

- 探索高能宇宙线起源并开展相关的高能辐射、天体演化、暗物质分布等基础科学的研究

实验装置:

- KM2A: 1.3平方公里的探测阵列, 5195个ED和1171个MD
- WCDA: 3个水池, 3120个PMT (900支8英寸, 2220支20英寸)
- WFCTA: 20台

现状:

- 2021年, LHAASO建设完成并开始全阵列科学运行
- 2023年, 通过国家验收

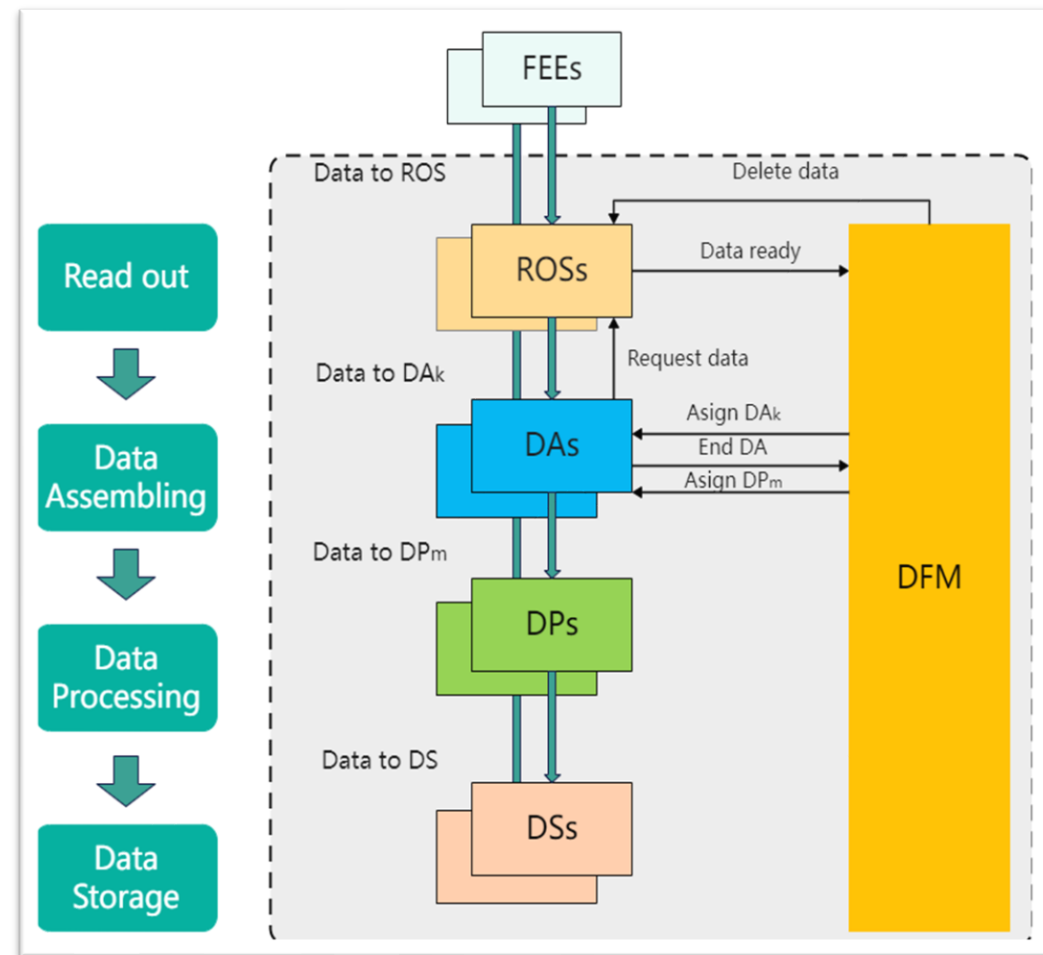
LHAASO DAQ现状与升级目标

■ LHAASO DAQ现状

- 在线数据处理采取流计算方案
- 数据流模块间耦合度高
- 修改在线触发算法困难

■ 升级目标

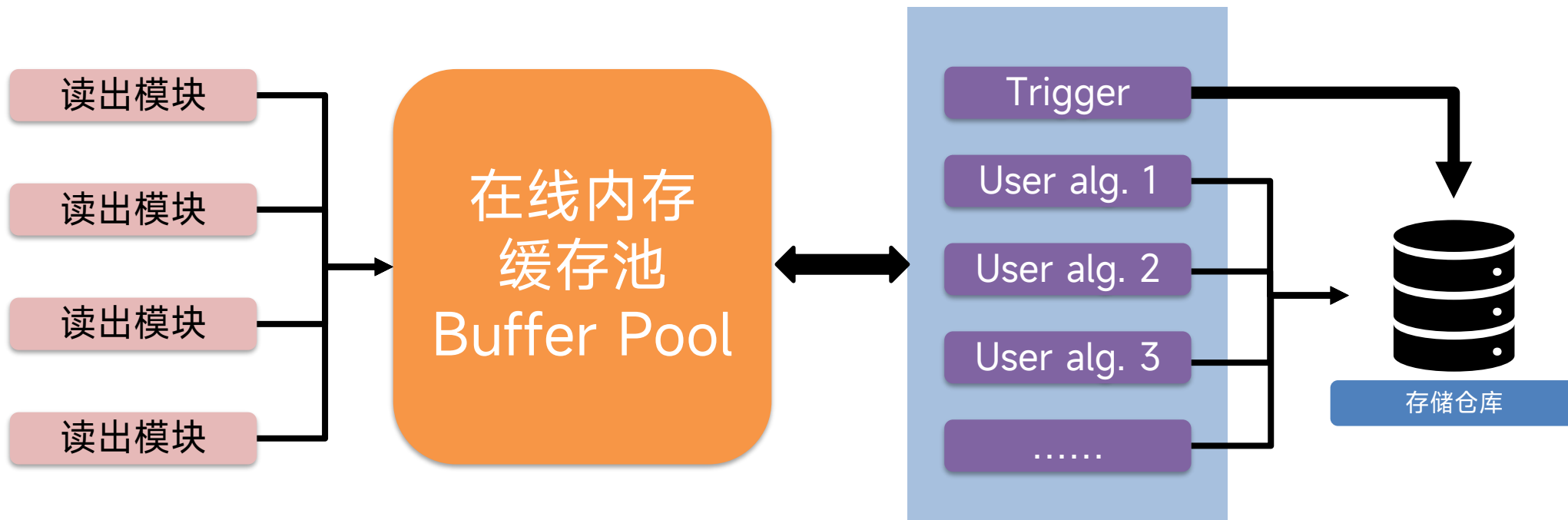
- 离线数据分析用户主导在线算法开发
- 为用户提供基于文件的在线数据处理软件开发方案
- 同时保证数据处理的实时性



LHAASO 现在采用的数据流架构

基于缓存的在线计算软件架构

OPF
在线处理框架

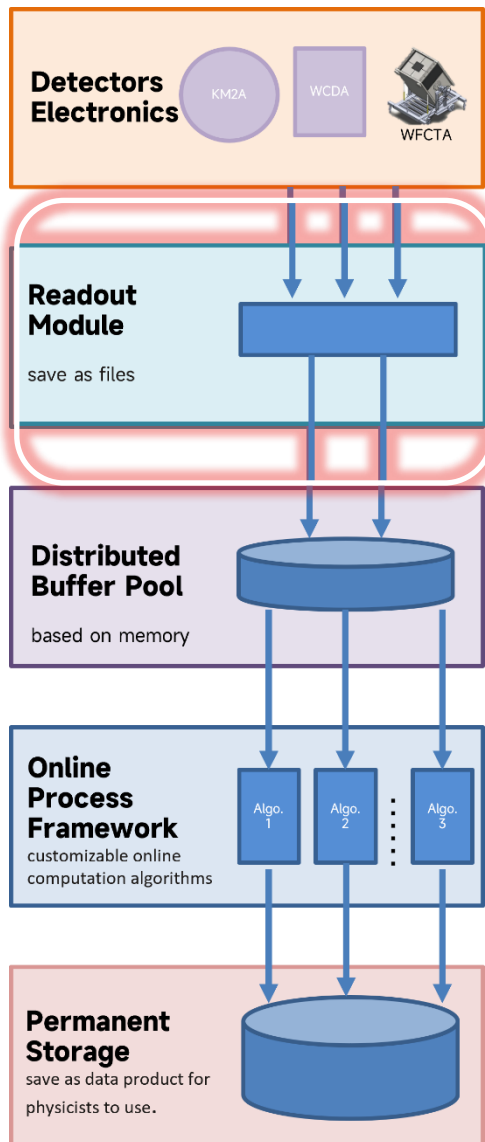


 文件形式流动

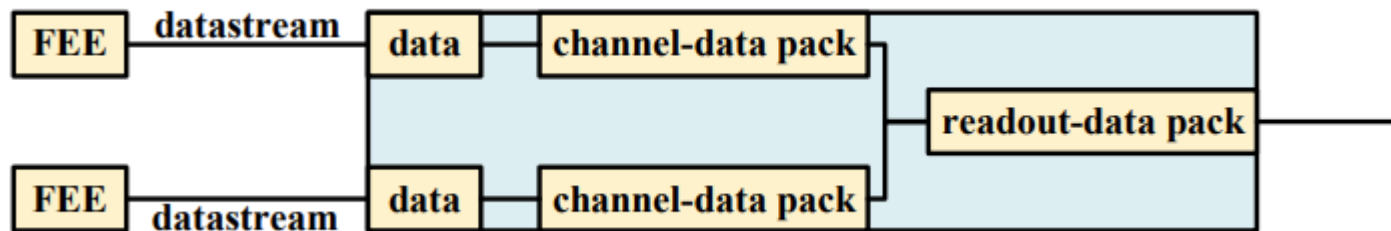
分布式内存作为缓存

支持多条在线计算

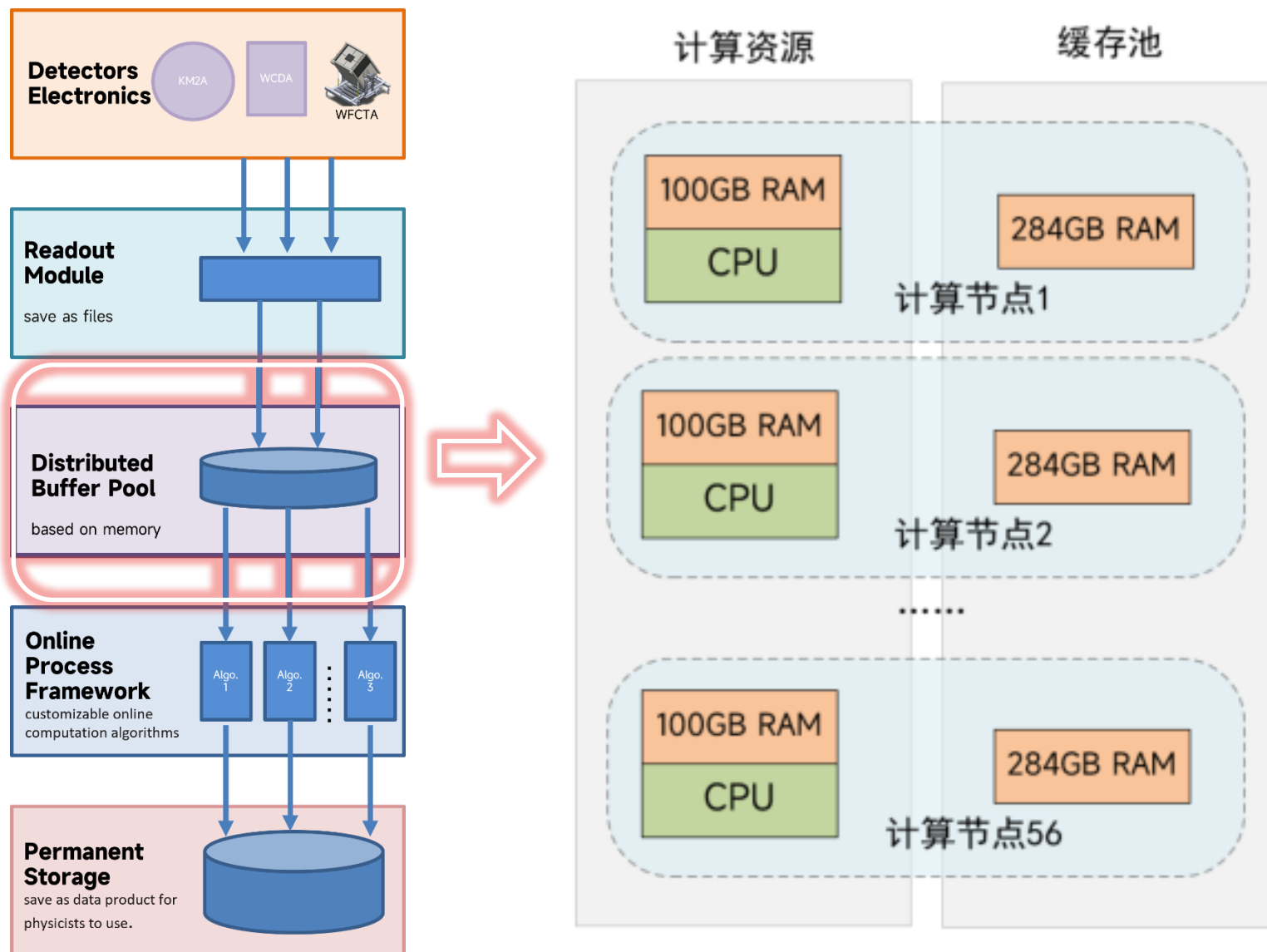
读出模块



- 对指定探测器与电子学进行配置
 - 从电子学获取数据
 - 对数据进行质量检查
 - 对数据进行切片标号
-
- 数据包组装，并写入缓存



LHAASO分布式内存缓存池



56台高密服务器

- 在线计算 (3000 CPU cores)
- 数据缓存 (16TB 缓存池)

为用户提供:

- 超过**30分钟**原始数据缓存
- 计算资源

LHAASO分布式内存缓存池——技术路线

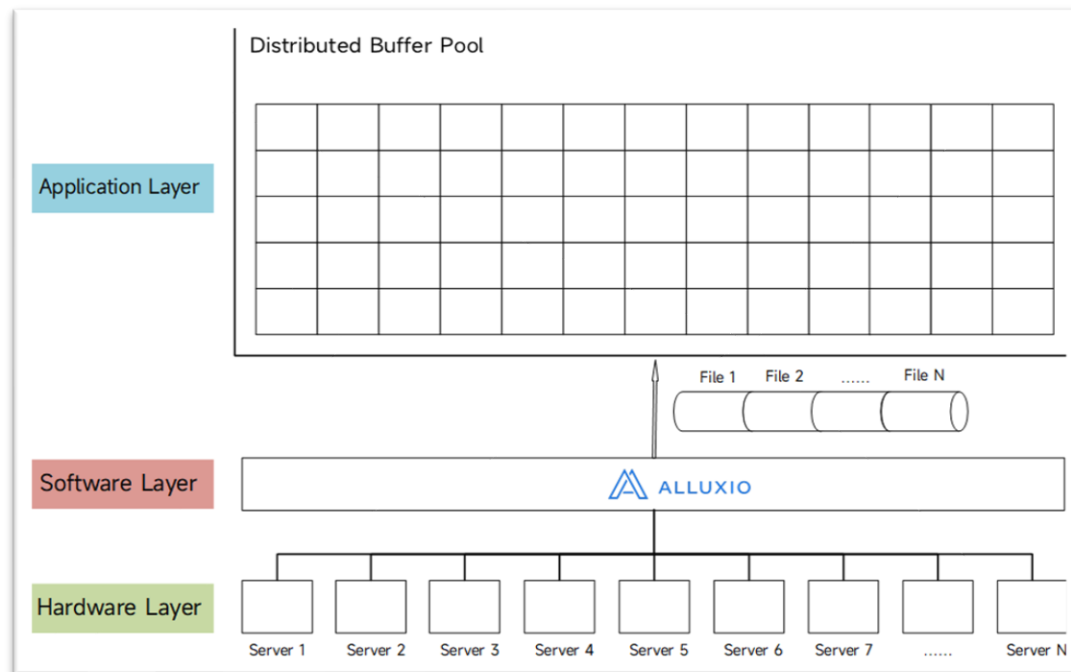
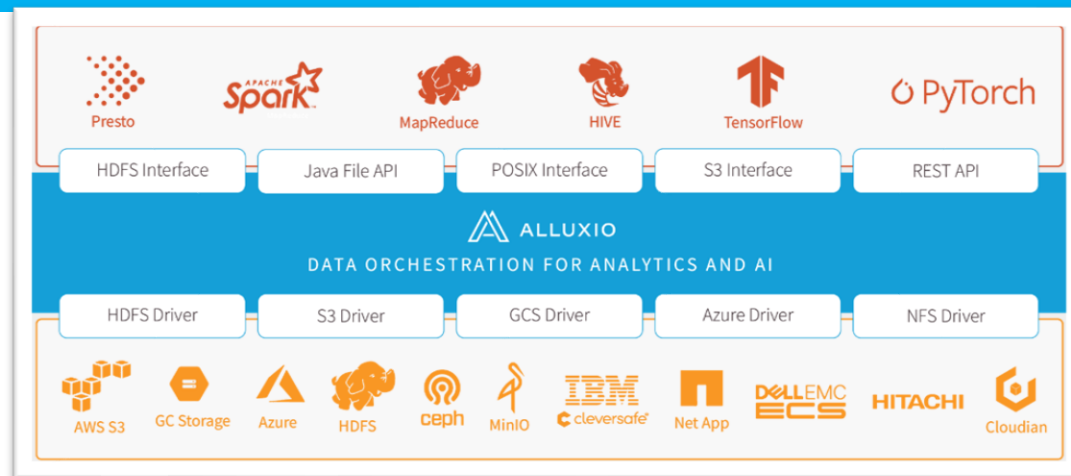
Alluxio

- 社区版基于 Apache 2.0 证书开源
- 分布式内存型数据编排平台



把复杂存储系统统一成单一数据源
跨存储系统抽象数据访问，虚拟化所有数据，
并通过具有全局命名空间的标准化 API 将数据
呈现给数据驱动的应用程序。同时，具有缓存
功能，以便快速访问暖数据。

- 管理分布式内存缓存池



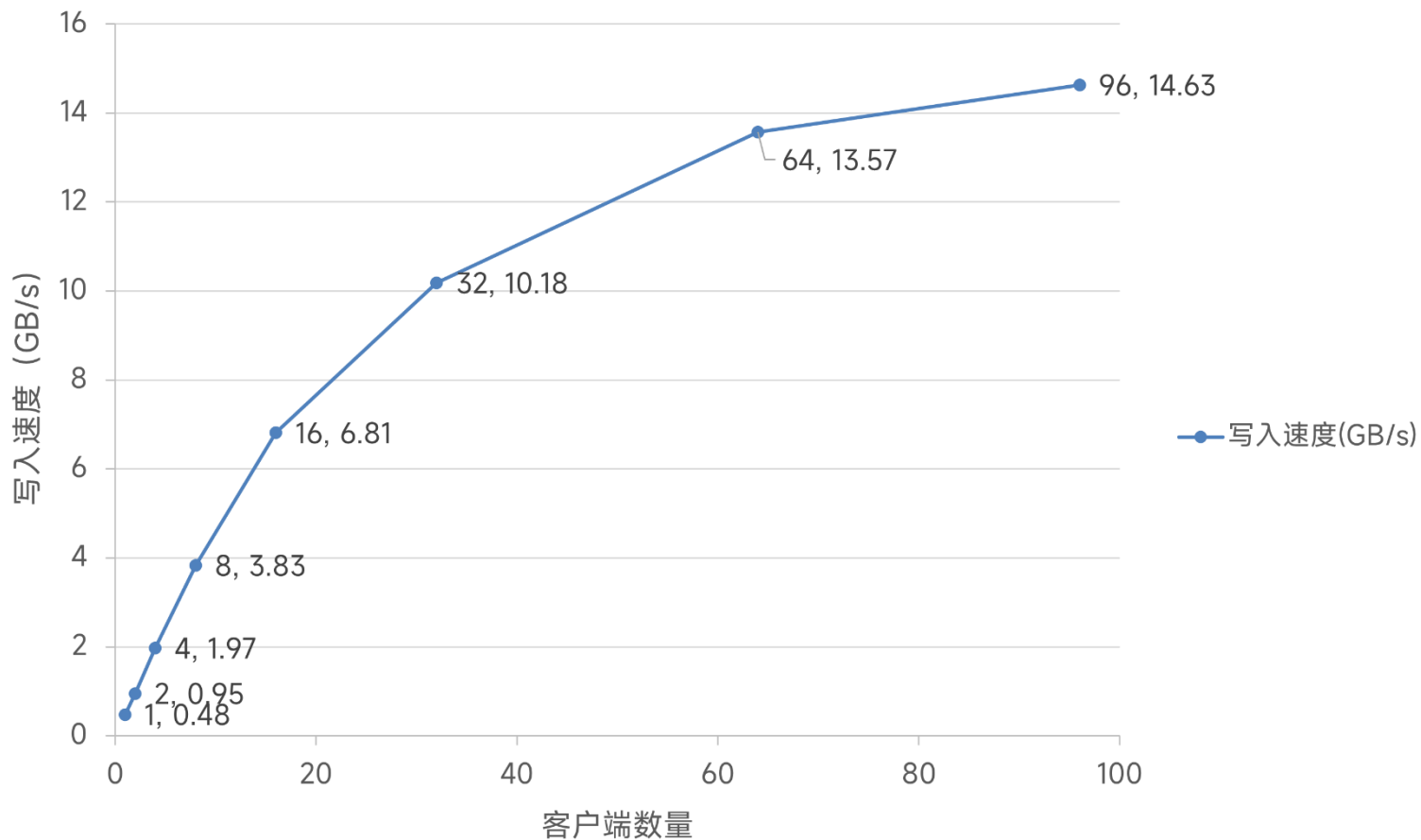
基于Alluxio的在线缓存池性能测试

测试条件

- CPU @ 2.70GHz (56 Cores)
- 384GB RAM
- 20个节点部署Alluxio集群
- 写入块大小10MB，文件大小10GB

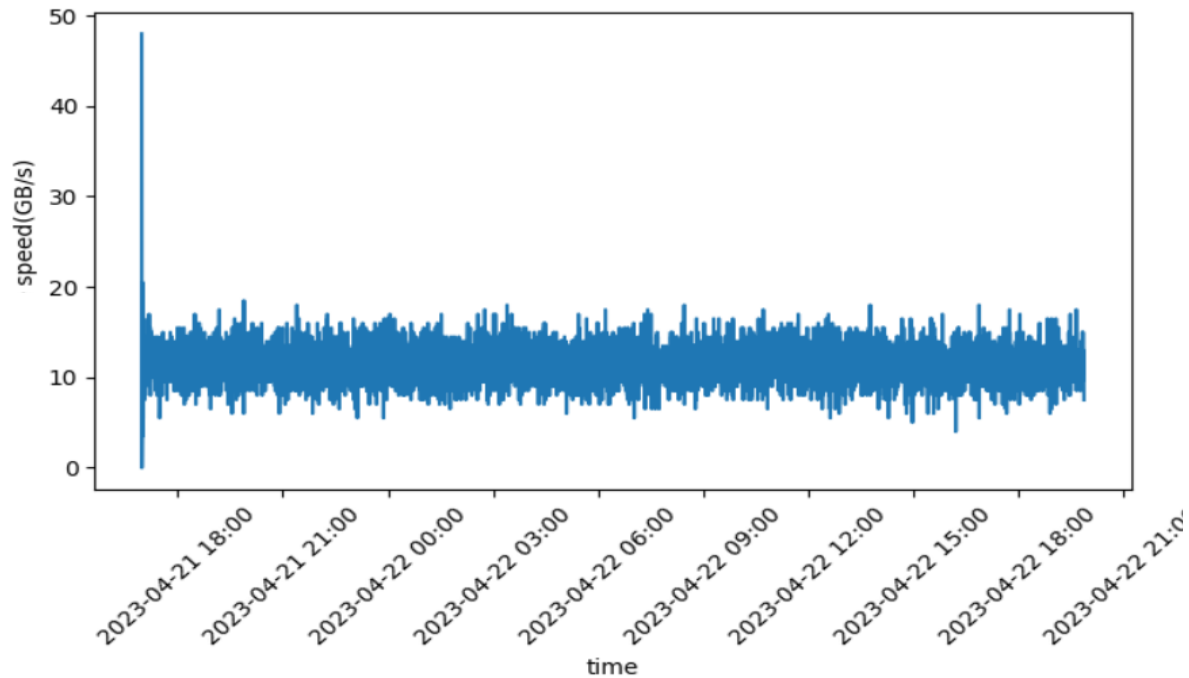
当前测试条件下，缓存池写入
性能为 **14GB/s**

客户端数量与写入速度的关系

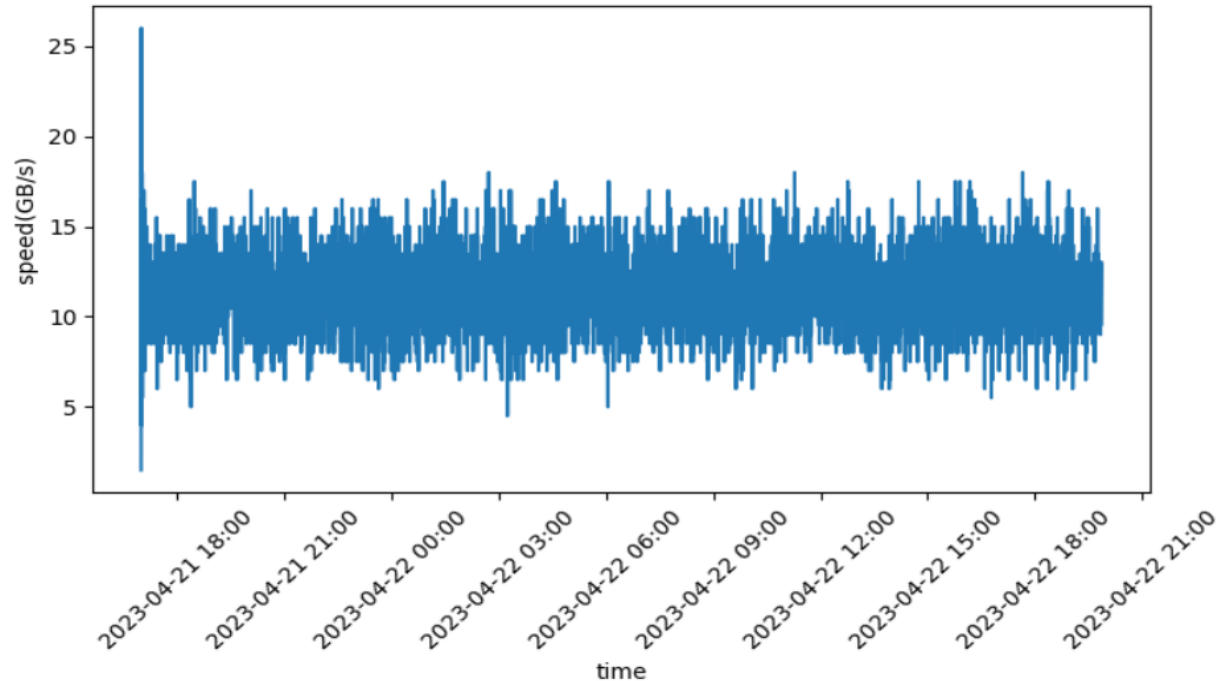


基于Alluxio的在线缓存池性能测试（读写混合测试）

缓存池写入性能

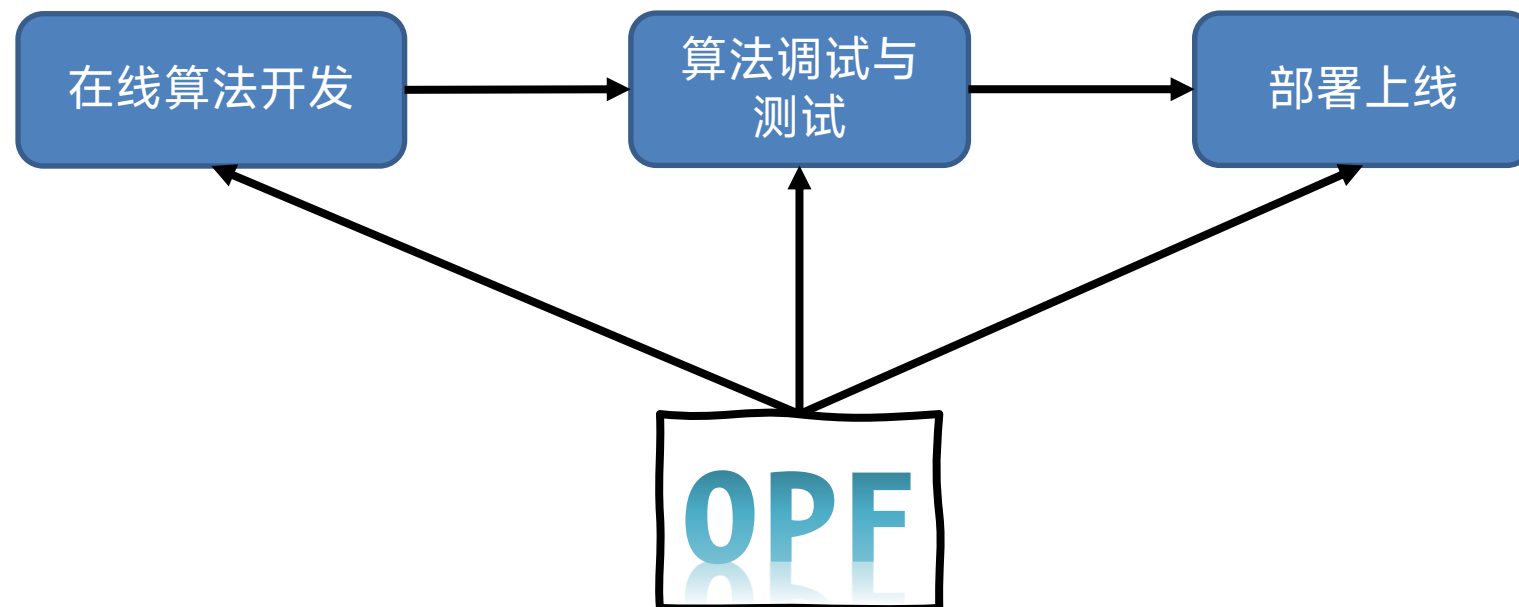
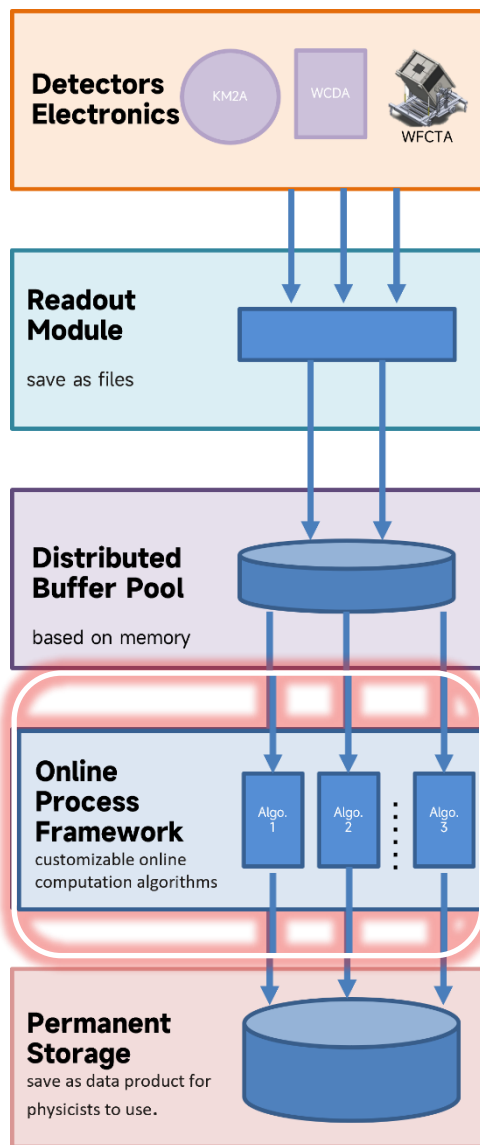


缓存池读取性能



经过超过24小时读写混合测试，缓存池稳定运行，平均读写总吞吐率超过**22GB/s**

Online Process Framework OPE



- 缓存池中数据访问规范
- 算法开发与调试
- 数据驱动、流处理
- 资源调度管理

WCDA与KM2A联合触发（基于新在线计算方案）

WCDA触发后获取KM2A原始数据

- WCDA触发得到输出事例时间段
- 基于在线缓存的KM2A原始数据，将WCDA事例时间段内的数据打包保存
- 数据解码，生成ROOT文件

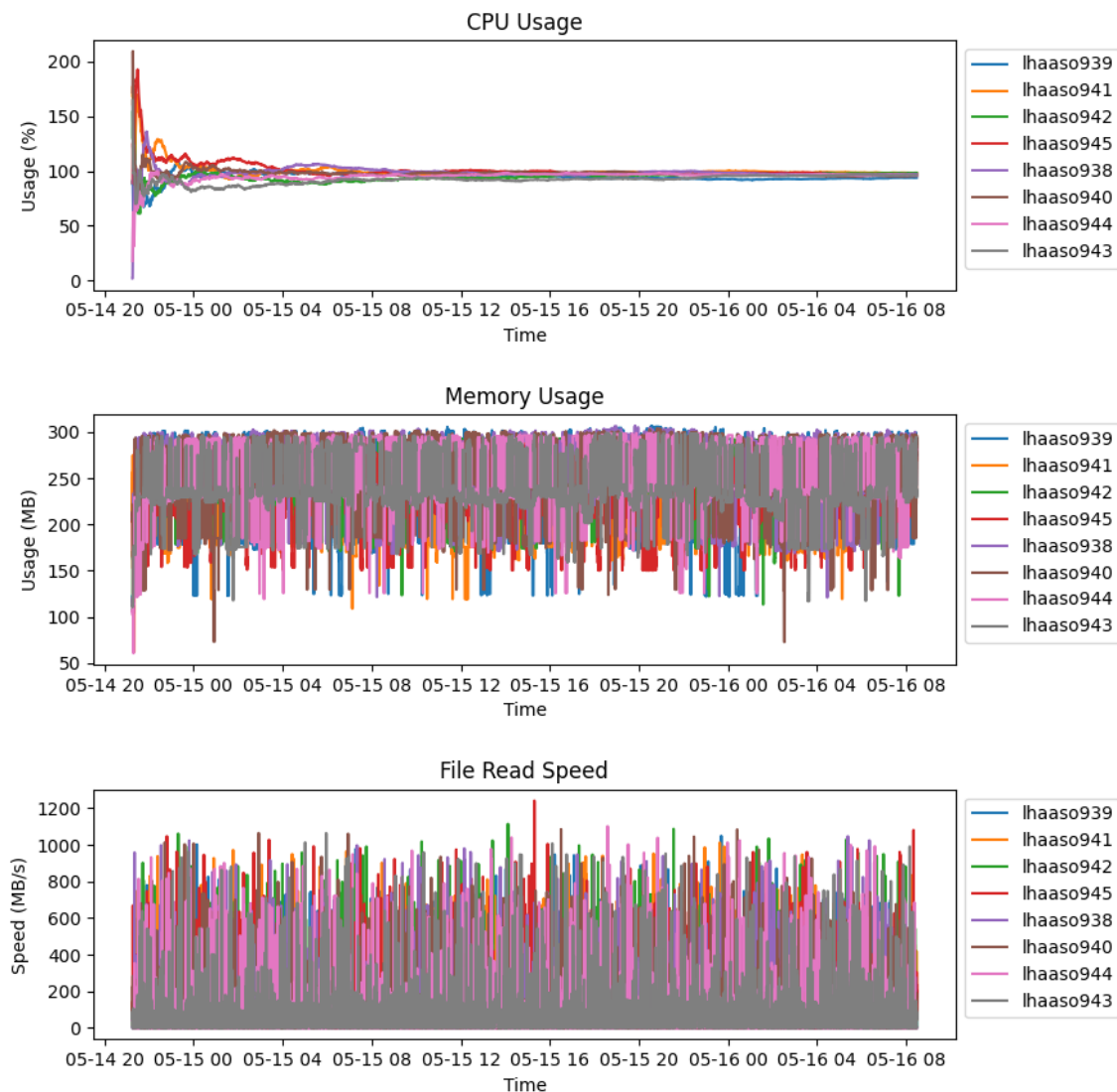
触发配置参数：

- 数据窗口：2us
- WCDA触发率：32KHz

存盘数据率估计

- KM2A事例格式保持一致，数据率：40.8MB/s

5月31日，开始正式运行





物理目标

- 在Scaler模式中，只记录在固定时间间隔内到达探测器的粒子个数。利用Scaler模式数据中计数率的变化，可监测探测器的运行状况，还可研究短标时的物理现象，如 γ 射线暴、太阳物理中的GLE事件和福布斯下降、雷暴期间宇宙线强度变化等。

所需数据

- 目前需要KM2A中ED数据，统计每一个探测单元在符合时间窗口为100 ns内多重数(ED个数)分别为 $\geq 1, 2, 3, 4$ 时的计数，每0.1 s记录一次。

存储信息

- 包含时间信息，以及每一个探测单元中4种不同多重数的计数。

6月20日，开始正式运行

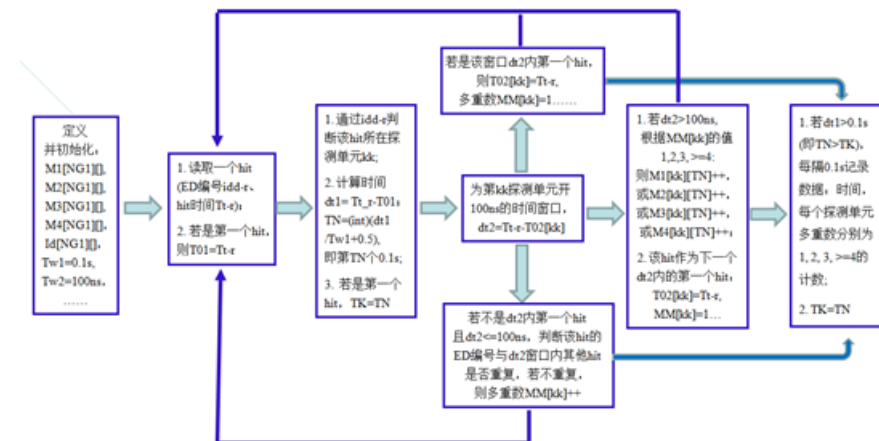


图3 Scaler 模式数据的数据处理流程图

数据处理流程图

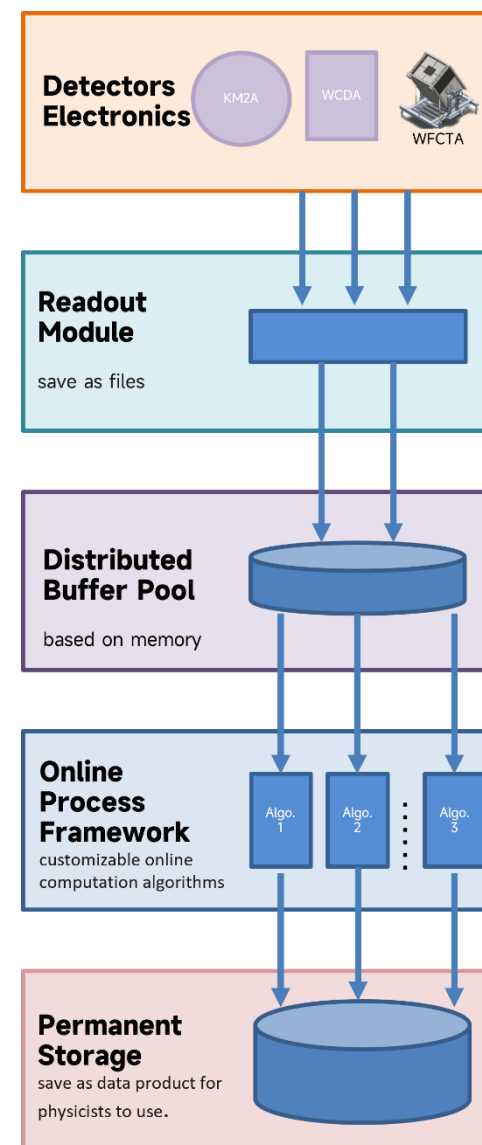
LHAASO在线计算软件升级（LHAASO DAQ Run II计划）

- 目标：研发新的在线计算软件，提高系统稳定性，拓展在线计算应用场景
- 任务：
 - Step 1 方案验证
 - Step 2 保留现有数据流，在新的模型下完成触发算法迁移
 - WCDA与KM2A联合触发 ✓
 - 单粒子模式触发 ✓
 - Step 3 LHAASO整体切换到基于在线缓存的计算模型



总结

- 设计了一种基于分布式内存缓存的在线计算框架，为用户提供基于文件的在线数据处理软件开发方案
- 该在线计算模型已在LHAASO实验中成功应用，实现WCDA与KM2A联合触发、LHAASO ED探测器单粒子触发模式

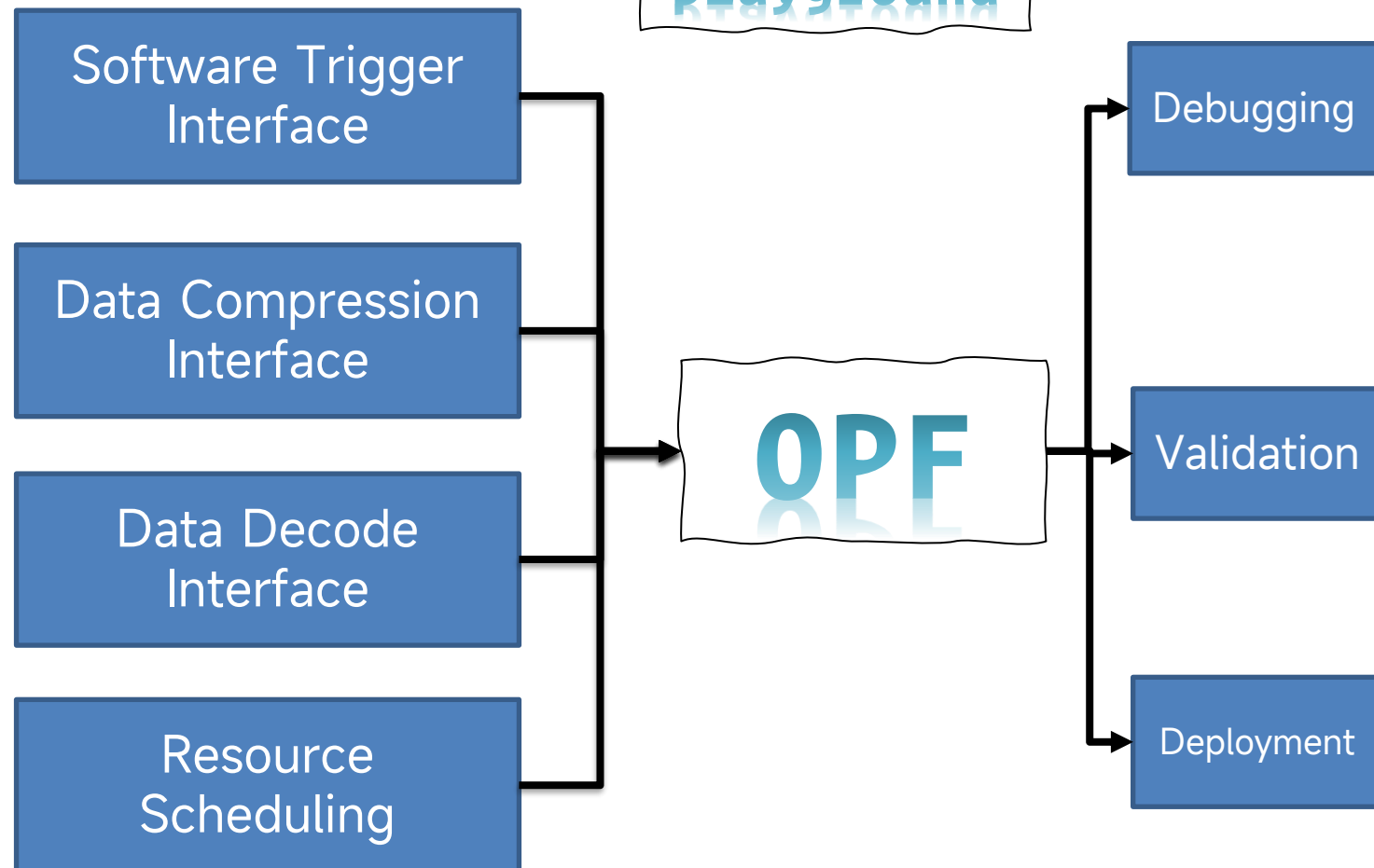


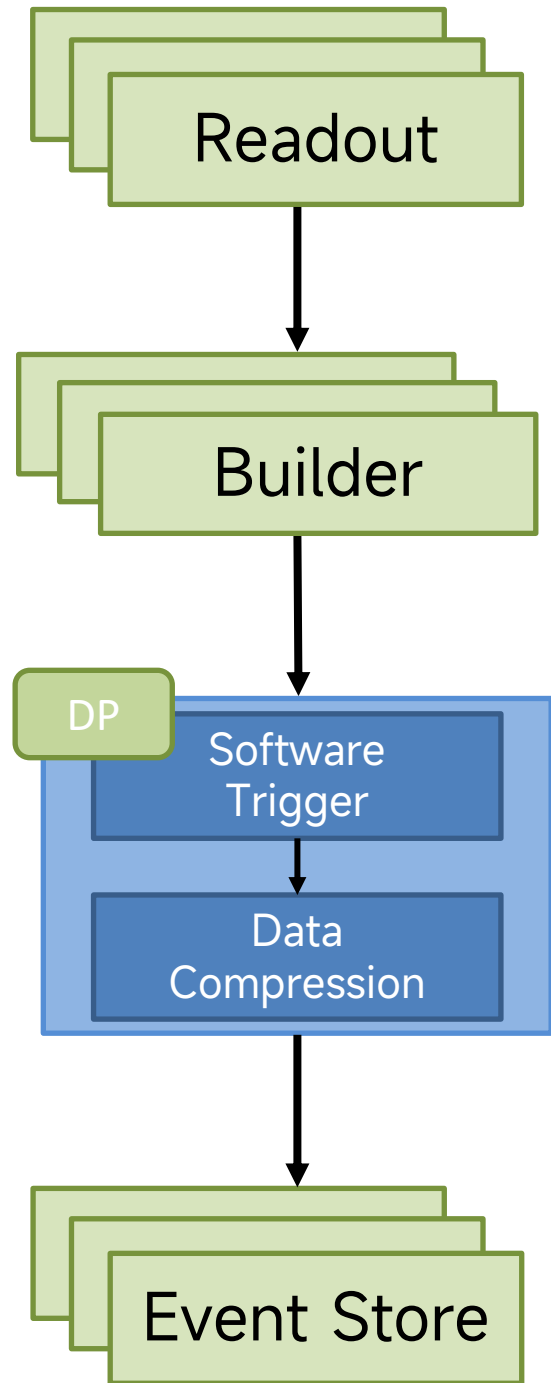
An aerial, grayscale photograph of a large, circular stadium or arena. The structure features multiple concentric tiers or seating areas, with a prominent central circular field. The image is positioned on the left side of the slide, partially overlapping the white background.

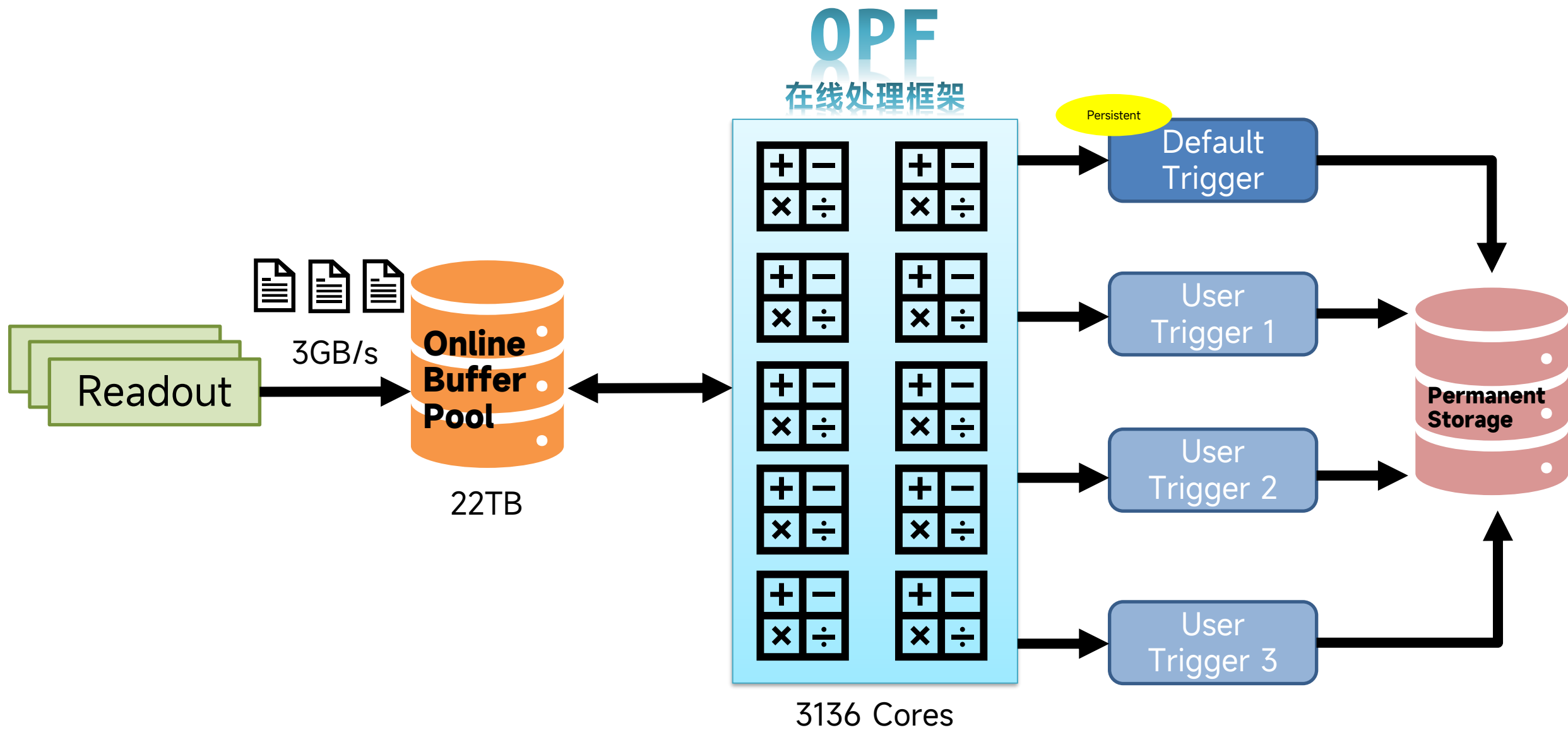
Thanks

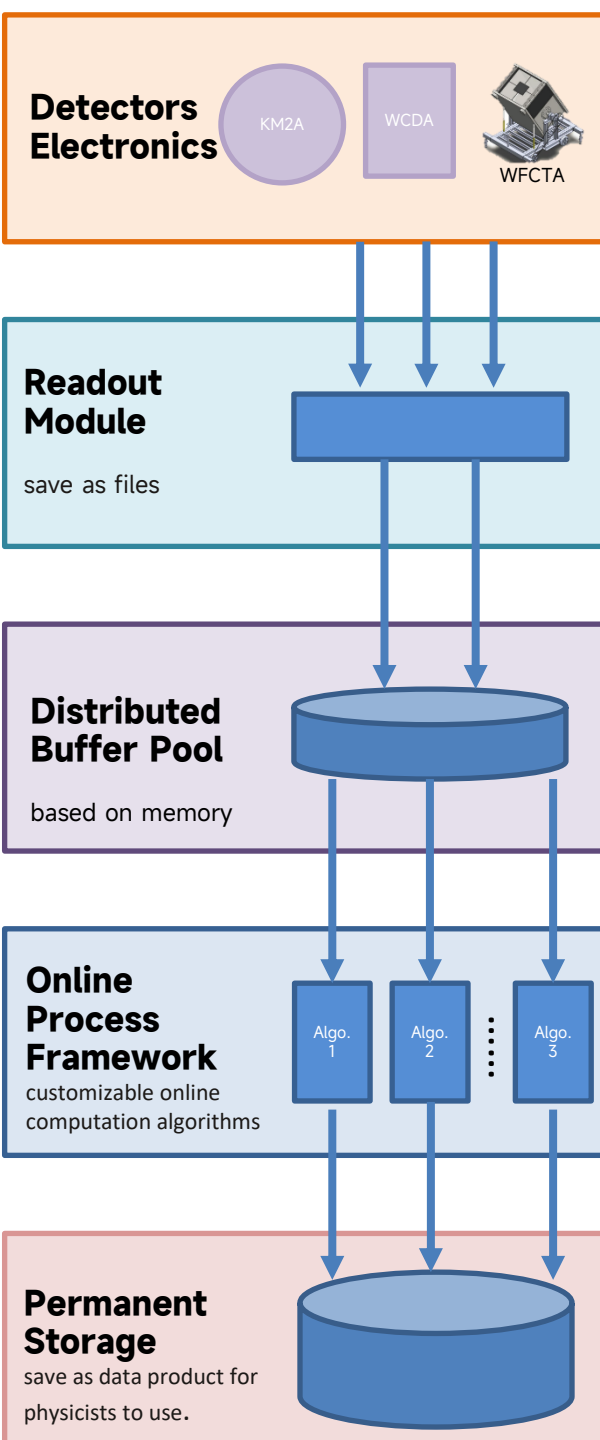
总结与展望

OPF
playground



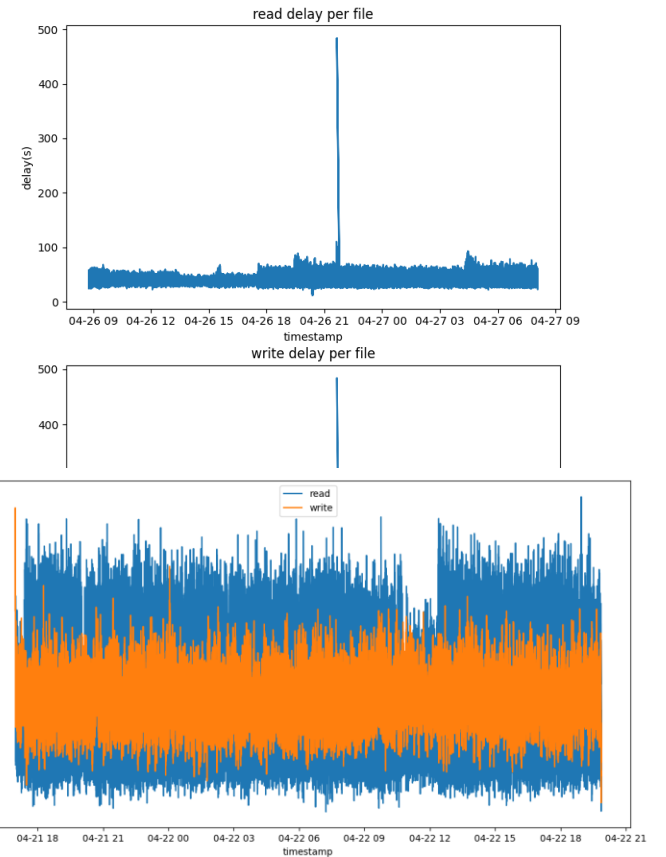






读写延时测试

24小时内仅6条记录读写用时大于100秒（共189047）



客户端数量	写入速度 (GB/s)
1	0.48
2	0.95
4	1.97
8	3.83
16	6.81
32	10.18
64	13.57
96	14.63

客户端数量与写入速度的关系

