

2024年度考核报告

邹佳恒

2024.11.20

岗位职责

- JUNO、BESIII和CEPC等实验的离线软件研发
- 时间安排（计划）如下
 - JUNO: 50%
 - BESIII: 20%
 - CEPC: 20%
 - 自由探索（量子计算）： 10%

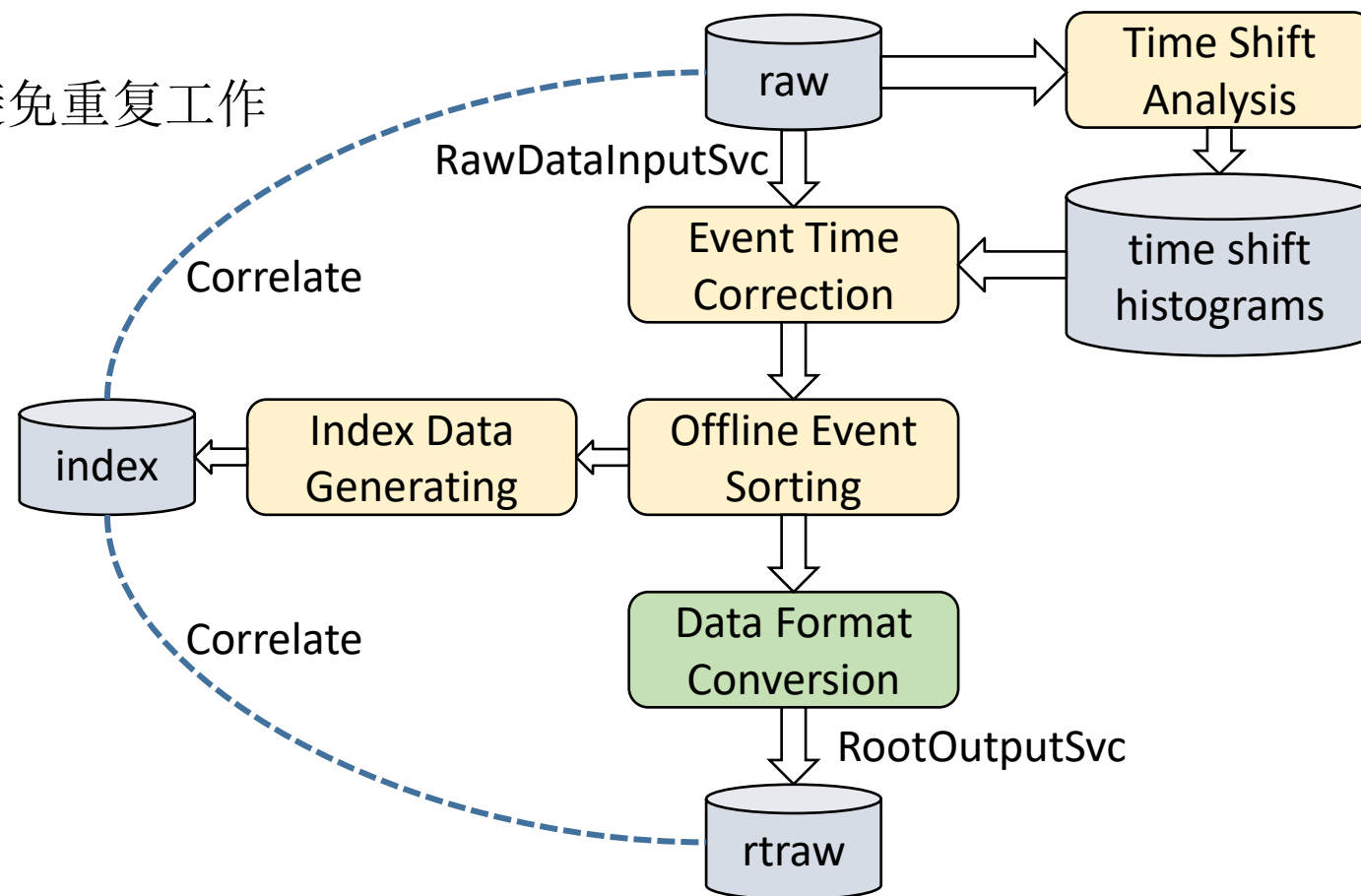
JUNO原始数据预处理全链条软件的开发推进

• 目的

- 完成离线排序等一次性任务，避免重复工作
- 方便后续重建等数据处理过程

• 主要内容

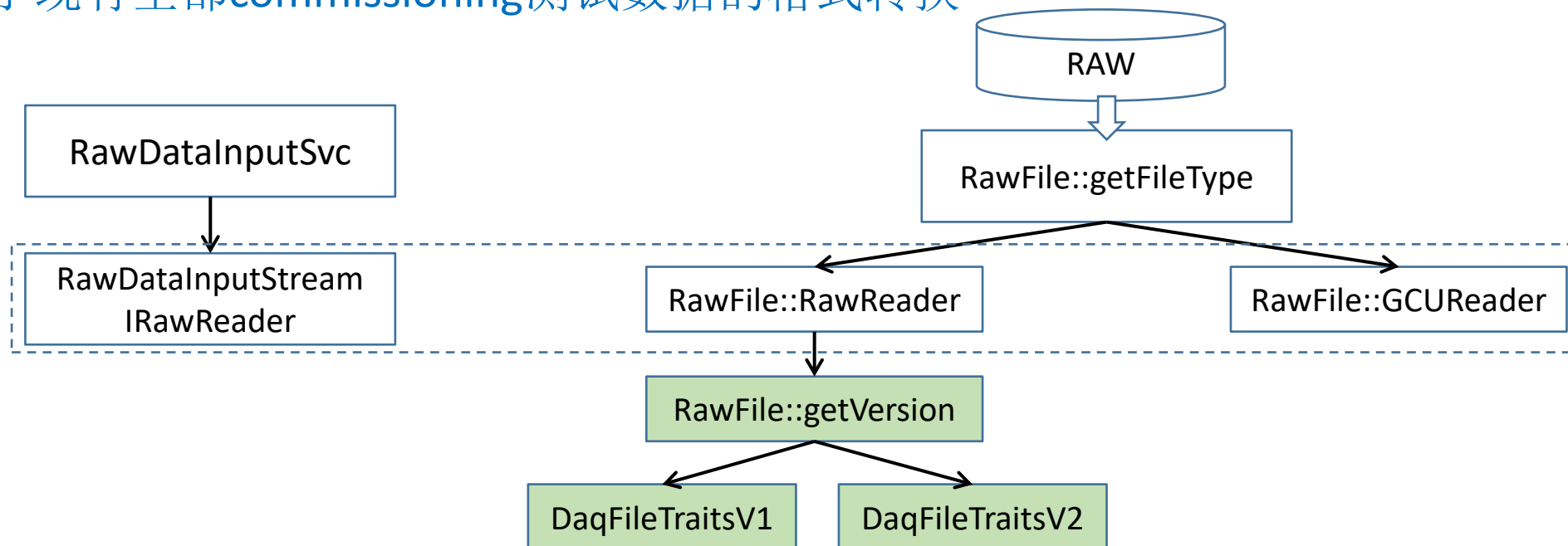
- RAW Data 读取服务
- 探测器时间对齐
 - 探测器时间偏移分析
 - 事例时间修正
- 离线事例排序
- 格式转换与事例索引



✓ 全链条软件的联合运行测试（有待根据真实数据情况完善优化）

原始数据读取与格式转换

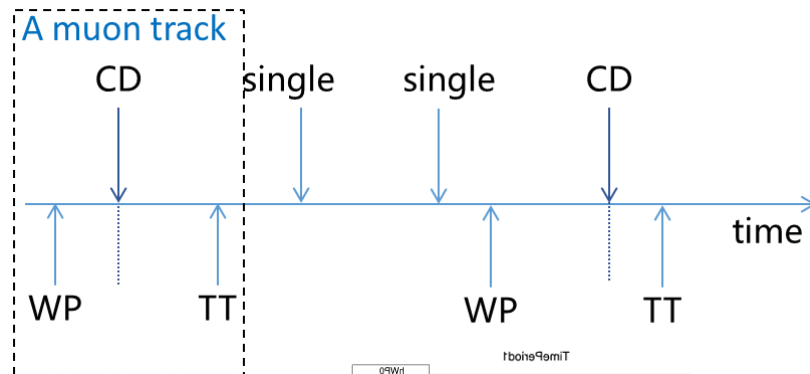
- 原始数据格式及版本（实验全生命周期）的兼容：GCU、DAQ组装后数据v1 & v2 ...
- 软件架构重构，便于升级维护：DaqFileTraits
 - 将不同数据格式版本的区别进行封装，软件自动识别，对用户透明
- OEC重建结果等新数据类型的支持，将根据实验测试进度持续扩展
- 已用于现有全部commissioning测试数据的格式转换



探测器时间偏移的模拟、分析与校准

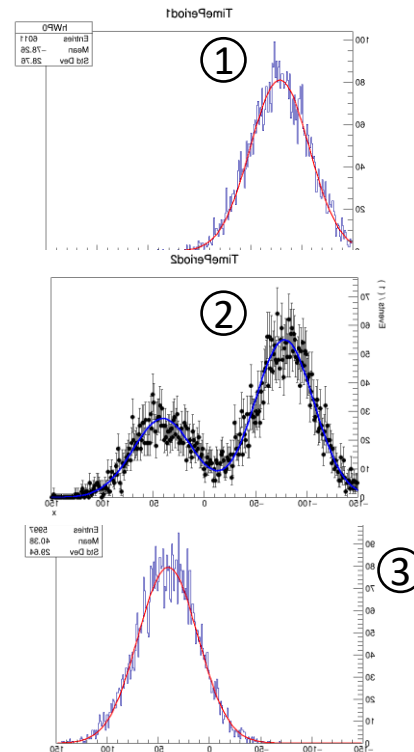
• 时间偏移的模拟

- Muon 径迹与 single 事例混合数据
- 仅考虑事例时间，且加入时间偏移量及其改变

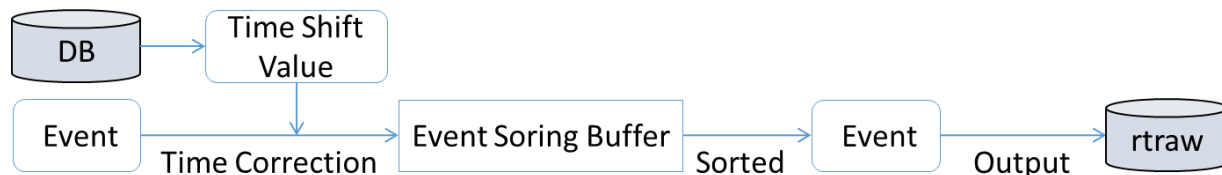


• 时间偏移的分析

- 周期性更新Muon径迹在不同探测器中触发时间的 Δt 直方图
 - 通过拟合得到该时间周期内的时间偏移值
 - 存入数据库（临时使用sqlite）用于后续的事例时间校准
- 周期内时间偏移发生变化的处理，在模拟数据上结果符合预期

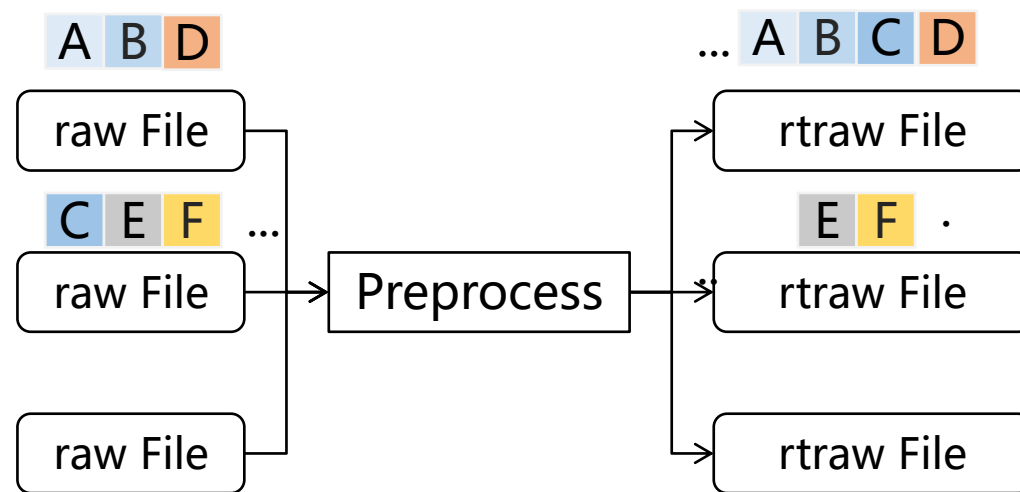


• 事例时间校准与离线排序



原始数据事例索引

- 中微子数据分析需要跨文件的事例时间精确排序
 - 离线事例时间校准与排序在文件边界处带来复杂性
 - rtraw与raw文件中事例对应关系被破坏
- 事例索引（index）数据
 - rtraw→raw的事例快速回溯
- 事例索引产生算法
 - 与事例排序算法联合运行

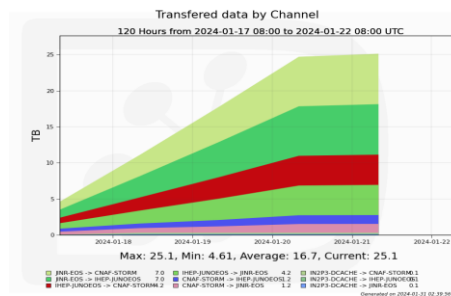


The events in raw and rtraw files do not correspond one-to-one

JUNOSW的多线程计算

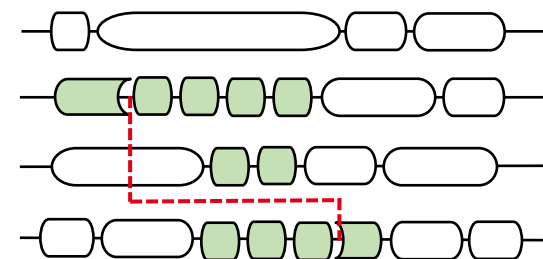
- 基于MT-SNiPER的事例级并行

- 年初通过了大规模的JUNO Data Challenge (DC-1)任务测试
- ~100k文件，4核作业，4个分布式站点的~3200 CPU核
- 国内大型高能物理离线数据处理软件多线程并行化的首次大规模成功应用



- 事例内并行

- 某些场景下事例级并行无法满足JUNO要求
 - 联合重建等算法需要串行执行
 - 在线事例分类（OEC）软件需要控制每个事例带来的延时
- 在SNiPER框架中实现了基于可中断微任务的并行调度功能
 - 与现有JUNOSW软件兼容性好
 - 已有串行算法移植难度低



波形重建算法的事例内并行化

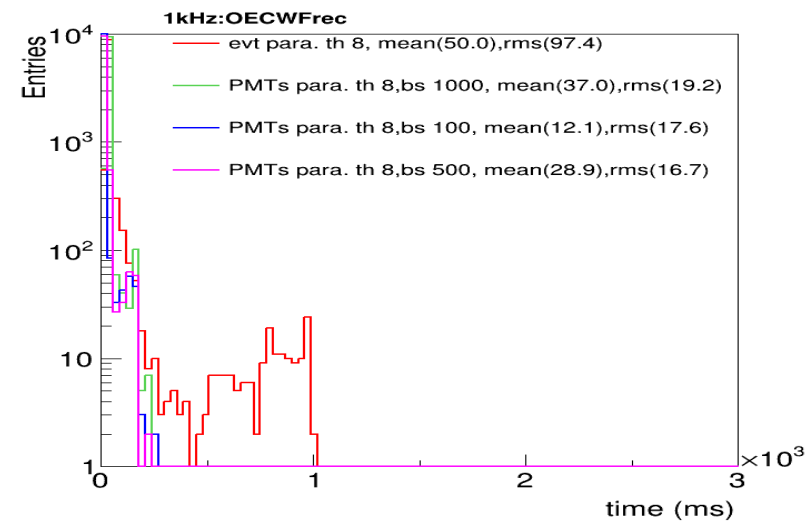
- 波形重建算法（OECWFRec）耗时较长，对OEC影响较大

- 将不同通道（PMT）波形分组并创建可并行的微任务
- 对微任务结果搜集合并

- 性能测试

- 物理结果与单线程软件完全一致
- 与事例级并行相比，大幅压缩了单事例的最大处理时间

- SNIper中新的事例内并行模型可靠性得到验证

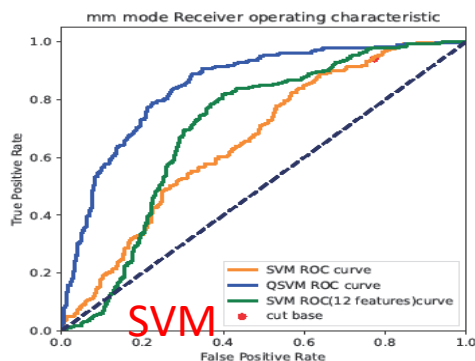
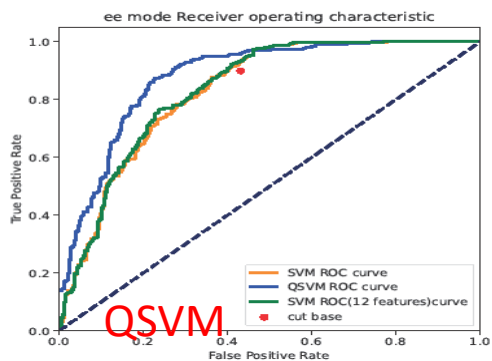


量子计算

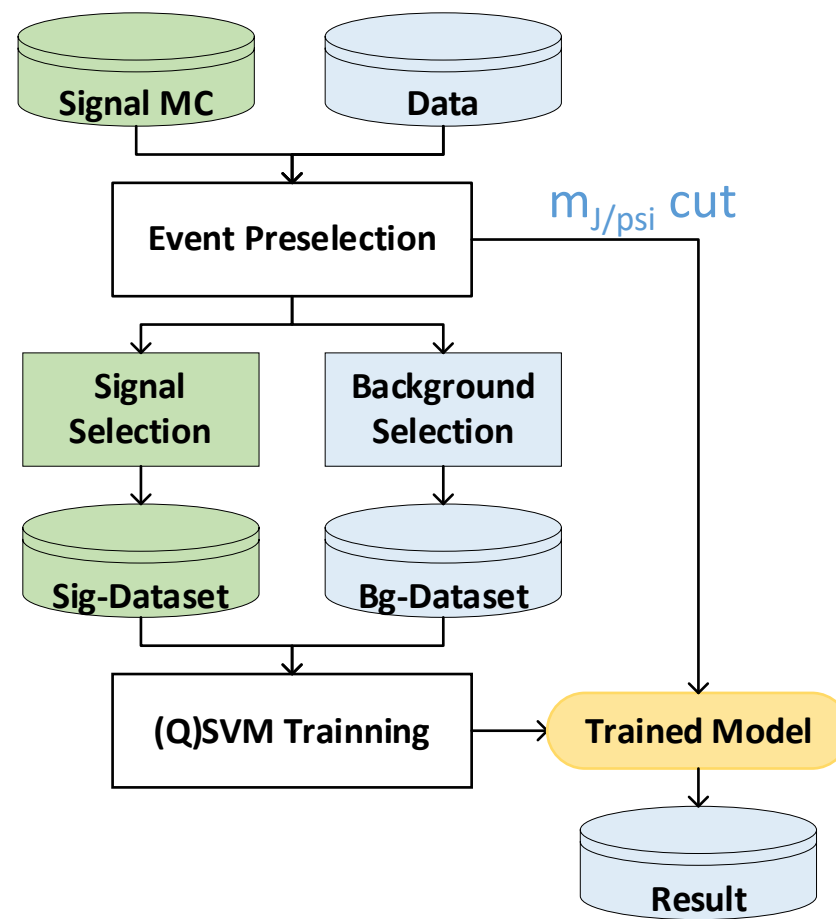
- 基于量子机器学习方法寻找Zc(3900)粒子

- 优化完善量子算法的应用步骤
- 训练数据集的准备
 - 保证MC与数据完全一致的事例预筛选条件
 - 优化本底、信号的选择条件
- QSVM与SVM的训练和优化标准

- 目前结果



- 优于cut-based方法的ROC曲线



其它软件工作

- JUNO OEC软件测试中解决了大量问题，提升了软件可靠性
- BESIII实验
 - 参与了BOSS系统升级工作
 - 操作系统以及大量外部库版本的大幅更新，Git、CMake等新工具
 - 系统升级策略与规划，疑难问题讨论等
 - CGEM数据格式的离线验证
- CEPC实验
 - 参与Key4HEP国际合作项目
 - 跟进Gaudi等底层软件的发展方向
- HERD实验采用SNiPER软件框架，参与相关软件研讨会

其它

• 主要研究成果与学术交流

- 一作期刊文章1篇： Offline data processing system of the BESIII experiment. *Eur. Phys. J. C* **84**, 937 (2024)
- 授权专利1项，排名第二： 分布式异构计算方法及系统， 发明专利号 ZL 2018 1 0163809.4
- 报告： 第一届高能物理计算用户研讨会， 基于SNIPEr的JUNO离线软件并行化

• **项目：** 参与重点项目申请一项， 但未批准（江门中微子实验数据处理与分析关键技术的研究， 中山大学）

• **毕业学生：** 硕士生1名， 联合培养硕士生1名， 博士生（副导师）1名

• 公共服务

- 参加软件组学生的季度考核
- 多篇核电子学与探测技术期刊文章审稿， 多名研究生毕业论文评审
- CEPC软件培训

下年度工作计划

- JUNO实验

- 推进离线软件的事例内并行化工作
- 推进在线OEC软件与事例内并行软件的架构整合
- 根据实验实际情况，实现对原始数据预处理软件的功能扩展和优化
- 为实验正式取数做好准备

- BESIII实验：软件系统升级，保证基础核心软件的可靠性

- CEPC实验：参与Key4HEP项目，完善CEPCSW的基础软件架构工作

- 与物理分析人员合作完善QSVM量子算法结果的合理性

Thanks!