

# 基于分布式数据流平台的 中子谱仪通用化数据处理框架

**CSNS Control Group**

**2023.07.09 青海·西宁**

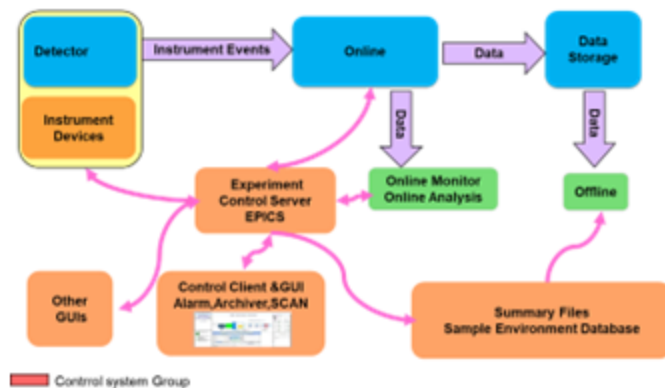
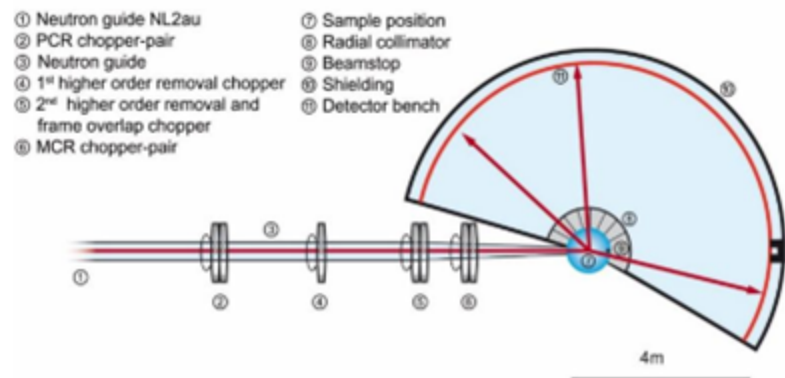
**第二十届全国科学计算与信息化会议**



## 大纲

- 背景与思路
- **DSNI设计目标和实现**
- 部署与验证
- 小结与展望

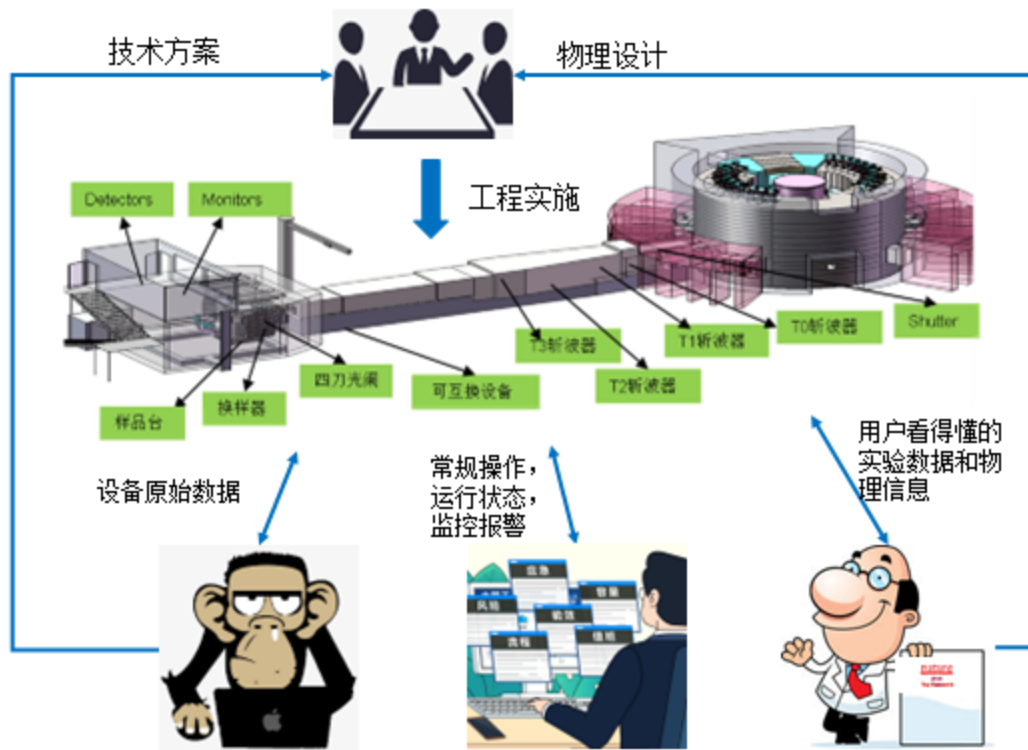
## 背景-中子谱仪装置的数据处理框架



数据处理链路：前端读出/控制通讯->在线分析/实验操作->离线存储

传统谱仪采用文件存储方式，导致闭环冗长，无法在实验级别实现所得即所见

## 谱仪控制与数据业务的需求分析



数据处理链条在每个环节归属于不同的使用者群体，他们对数据内容关注层面不同，需求偏重也完全不同。



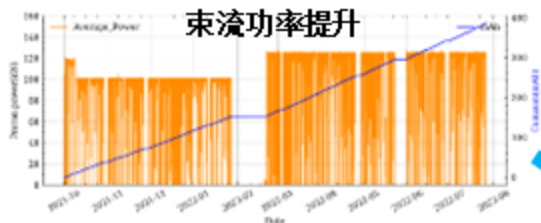
传统中子谱仪的数据处理基于文件传输：  
1. 传输效率低，不具备实时处理特性  
2. 接口复杂，系统耦合程度高

最终造成不同层面的数据业务糅杂在一起，对每一类用户都带来了复杂和不便

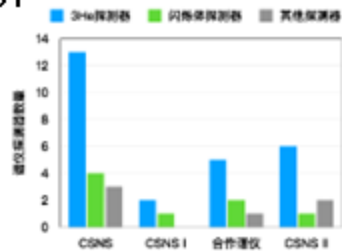
专业组：  
设备底层细节，原始数据检查，  
配置参数调节。。。

运维：  
监测直观，操作便捷，运  
作规范。。。

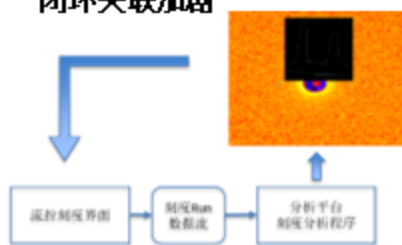
科学家：  
甲方用户，编排实验，设计方法，  
分析物理，focus on physics ©



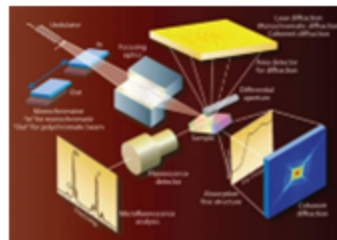
探测器规模、类别、效率、覆盖区域提升



实验方法多样化，与数据闭环关联加剧



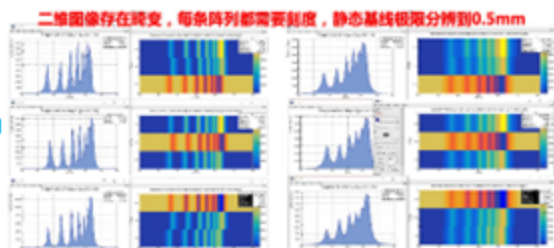
多模实验数据来源、类别、特征多样化，大数据趋势明显



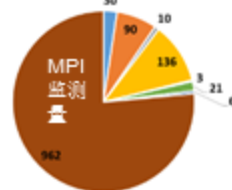
下一代谱仪装置提出的新挑战：

1. 更庞大的测量系统和数据流规模
2. 更复杂的实验方法：多模态、弱信号、准实时分析和反馈
3. 设备复杂度和部署运营成本

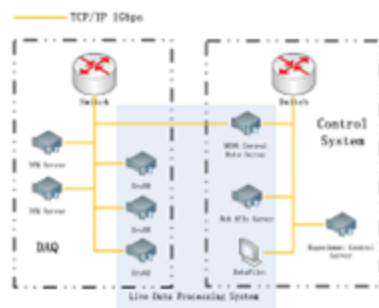
调试增多，复杂度提高，多设备协调比重大



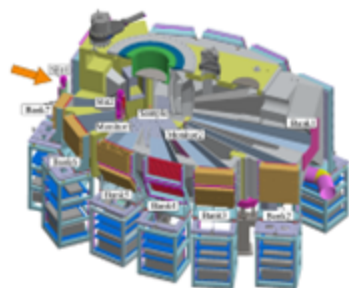
设备类型多样，部署管理复杂，监测量增多



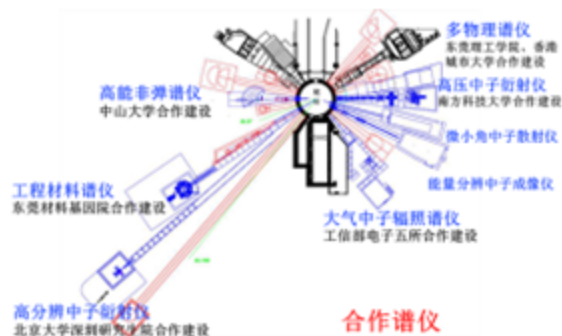
## 谱仪控制与数据处理演进方向



第一代解决有无：  
文件传输，实验可做



第二代解决性能：  
数据流传输，高效可靠，  
高吞吐高并发



第三代应该是迈向好用：  
面向用户，面向实验，承  
接物理，完备功能

## DSNI设计目标



标准化：统一模板，通用框架，标准接口



高性能：高吞吐，高可靠，准实时，可扩展

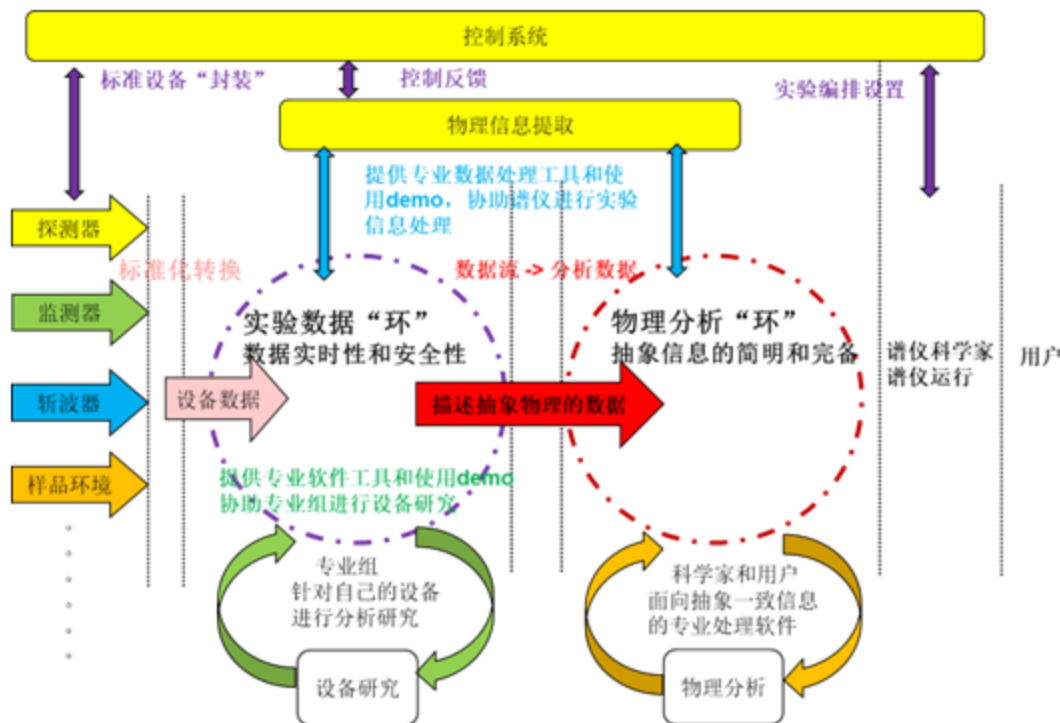


可编排：设备抽象，预处理分析，反馈闭环，模式编辑...



高可用：周边服务、设备管理、共享生态...

## 面向用户重新设计整体构架

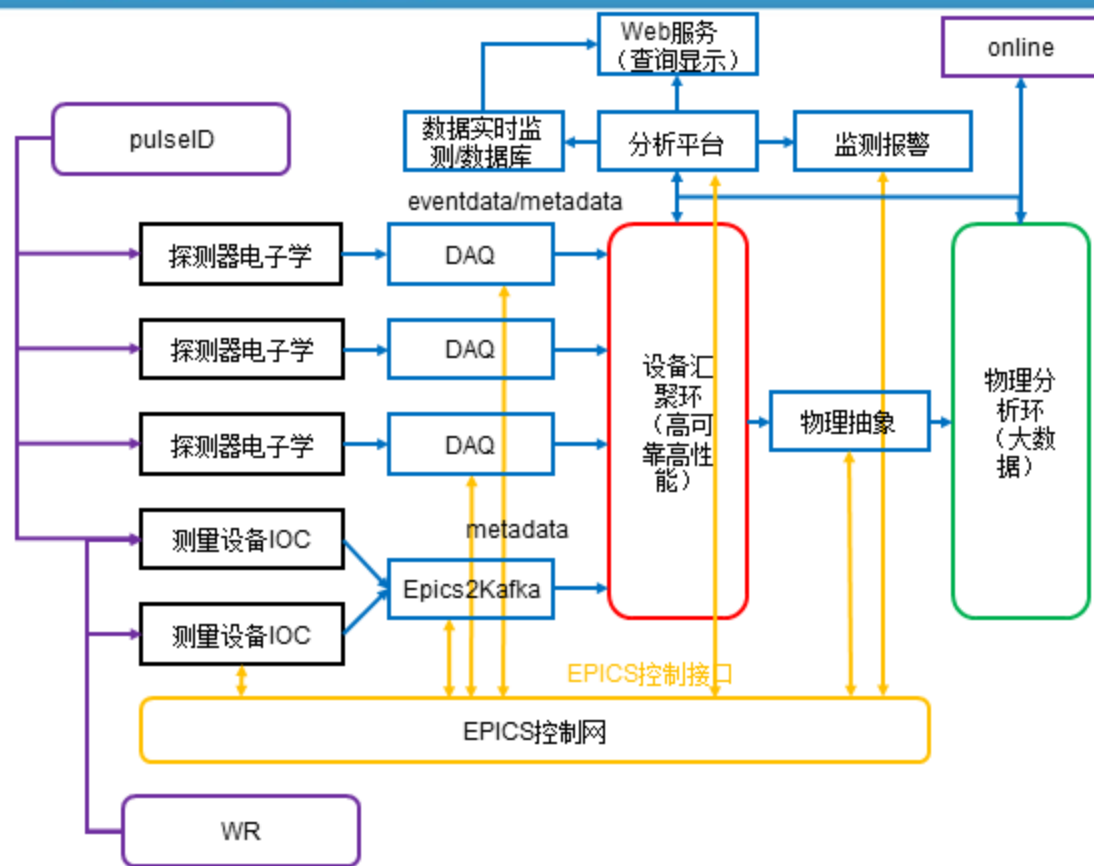


- 专业组关注设备数据交互
- 控制系统同时面向设备和实验物理，衔接双方
- 谱仪科学家关心物理信息交互，在物理层面控制实验



- 标准接口，大数据共享
- 控制抽象，数据处理的反馈和闭环
- 实时性，所见即所得

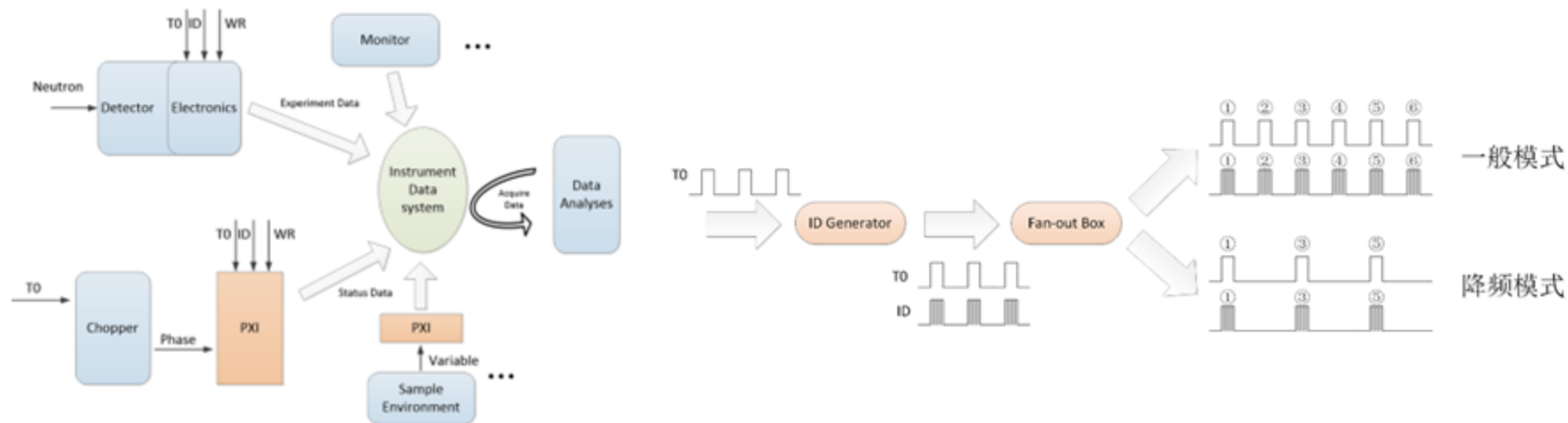




### 物理到工程实现数据组装融合

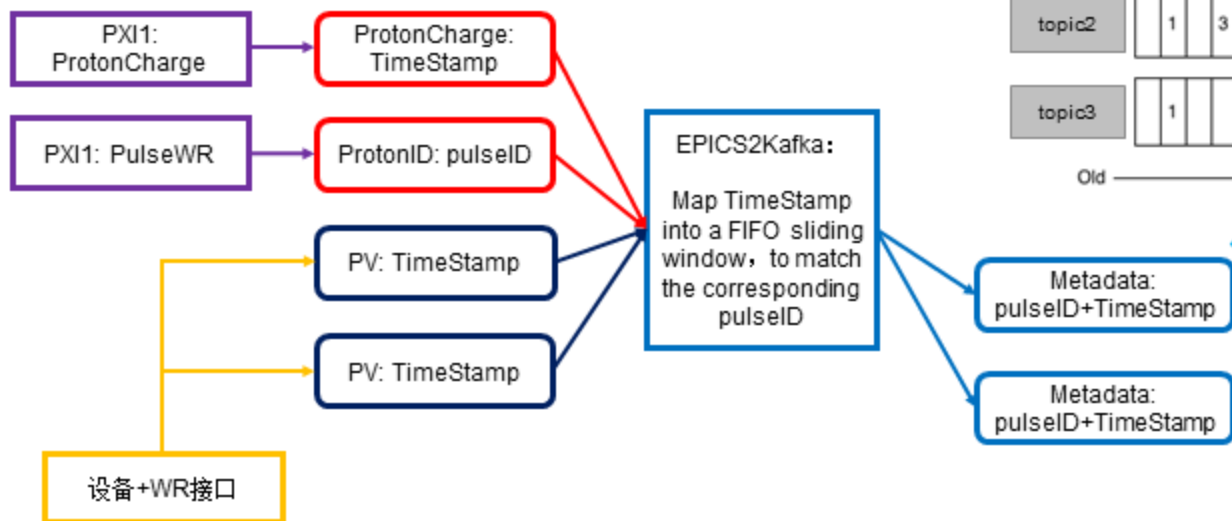
- 2种同步方式: TimeStamp和pulseID
- 数据获取框架: 实现分布式和可扩展
- 分布式消息处理平台: 实现数据标准化和共享
- 数据流分析平台: 可扩展和实时反馈
- 谱仪实验控制系统: 实时闭环
- 数据实时监测: 开放逻辑
- 报警系统: 多路分级, 智能扩展
- 其他周边子系统: 完善服务功能

## 分布式同步标签pulseID

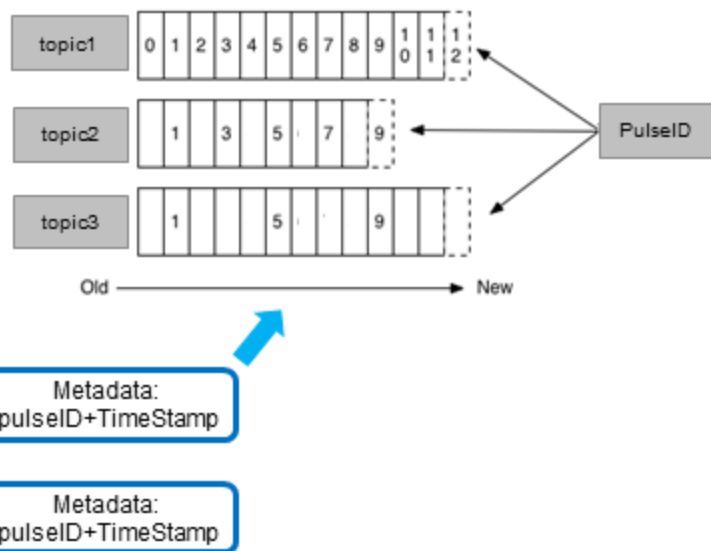


为单纯的TO触发增加带有同步编号的pulseID，给数据打上对应TO的数值标签，精确跟踪每一发TO的数据，同步聚合更加便利

## 高精度授时网络

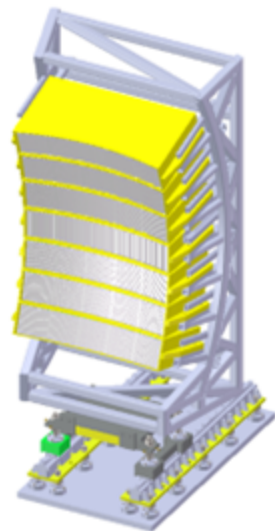


## Anatomy of a Topic

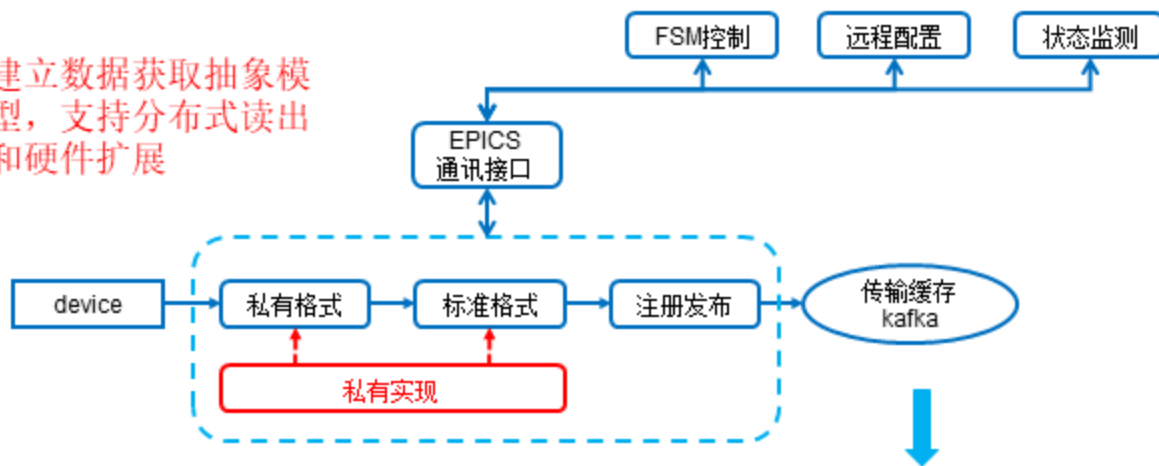


对不支持ID接口的第三方设备，继续提供基于WR的高精度时间戳（20ps），并在转化为标准接口的环节重新映射到pulseID，实现数据流同步标签的一致和对齐

## 基于消息流的高性能数据获取框架



建立数据获取抽象模型，支持分布式读出和硬件扩展



用例：利用metadata数据流发布中子重建统计信息，对设备运作情况进行长期监测和定位稀有故障

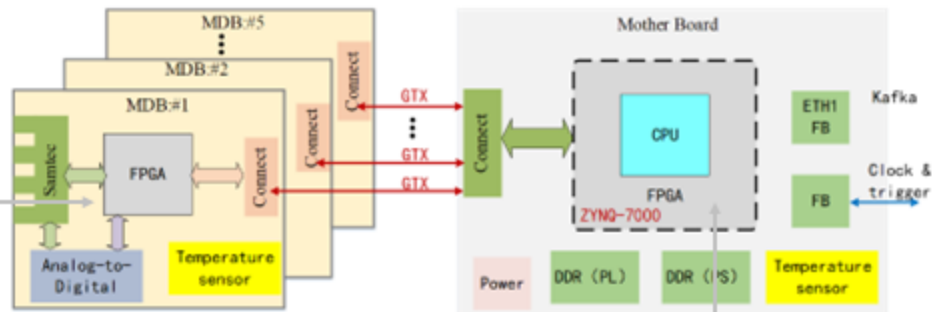


## SOC标准化电子学平台



特异性前端电子学

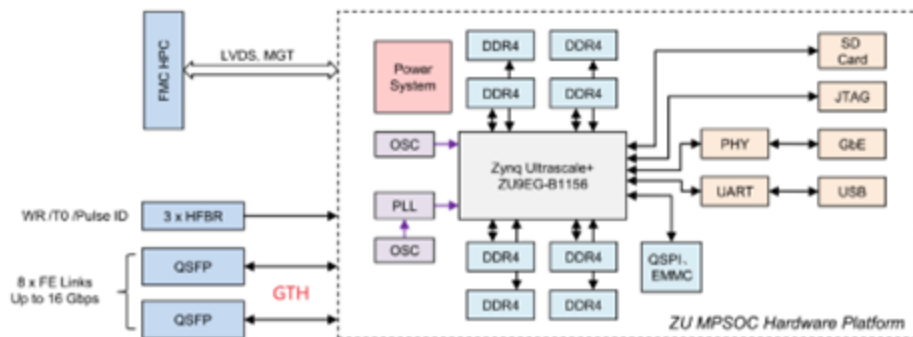
- CSNS工程材料谱仪
- 读出大面积闪烁体探测器



通用SOC电子学

- 通过光纤汇聚前端数据
- 基于SOC (FPGA + ARM)
- 实时数据处理, 通过kafka上传

配合通用化DAQ框架, 进一步设计了标准化电子学平台, 通过前端模块适配不同探测器类型, 实现硬件标准化

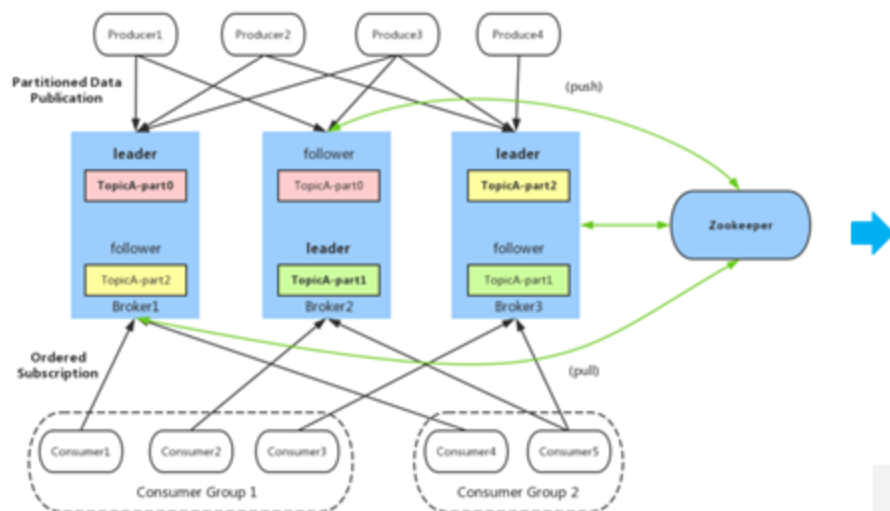


### 新一代SOC电子学平台

- 基于最新Xilinx Zynq Ultrascale+ FPGA
- 更高速的光链路：16通道16 Gbps 光链路
- 更快的片上CPU：4核Cortex A53 + 2核R5
- 更大的缓存空间：基于DDR4，最高可支持32GB
- 更高速的网络接口：10GbE/40GbE

前端带宽达到**256Gbps**，后端网口输出  
**40Gbps**

## 分布式流处理平台——选择与优化



Kafka平台具备高效可靠的特性，但目前仍受磁盘随机读取瓶颈的潜在制约，需要针对谱仪数据流特征适配优化。

数据流接口：随机offset寻址和延迟算法

避免频繁的IO抢占和传输阻塞

Kafka设置：收发缓存，工作线程，分区排序算法与封装

适配数据流特征，提升并发性

系统设置：JVM缓存，系统页缓存，网络缓冲区

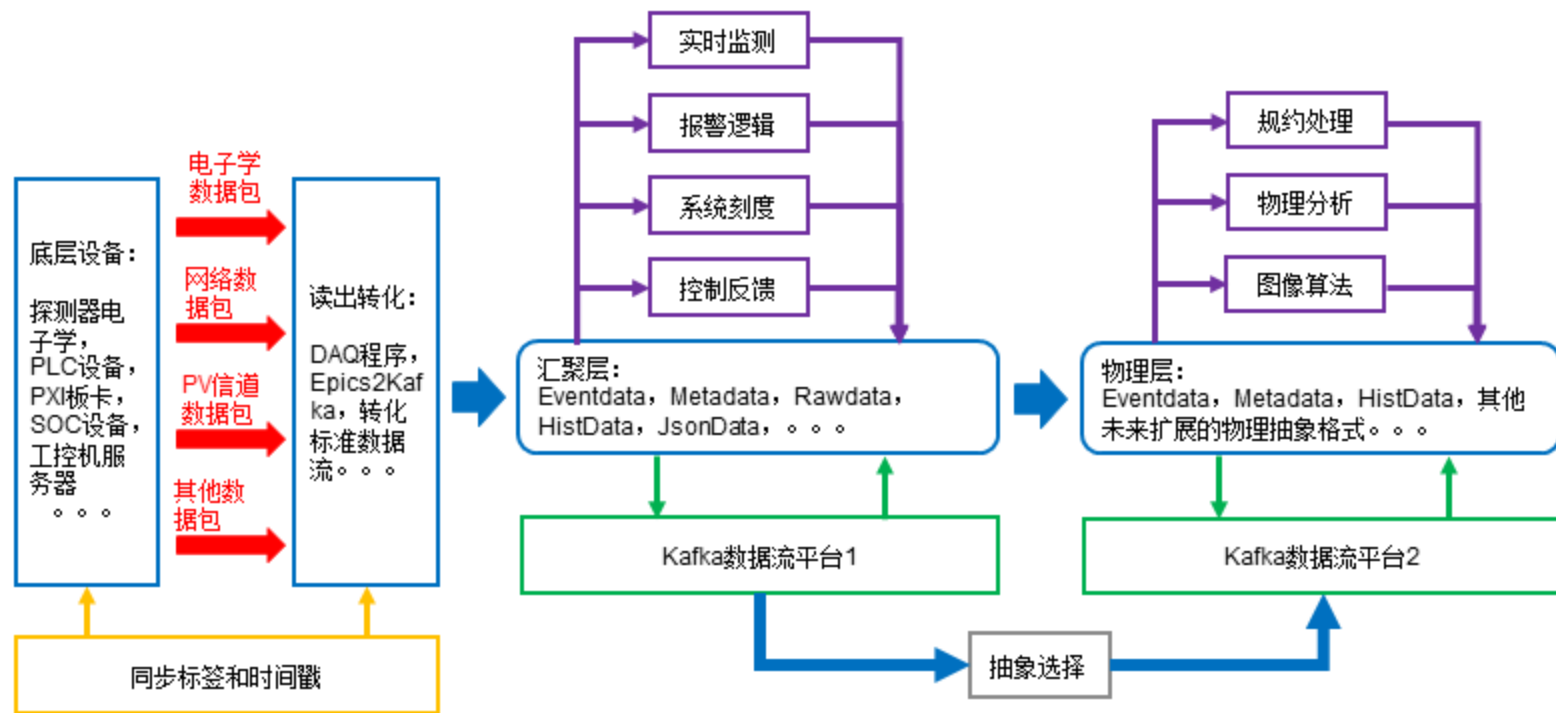
充分利用系统资源，优化网络传输

文件IO：预读机制，存储架构（未来）

降低随机IO比重和缺页延迟

	数据回溯	冗余备份	扩展性能	并发性
传统数据传输	无，小容量	各路之间相互隔离	强耦合，需设置	高
Kafka分布式消息平台	有，大容量（基于磁盘）	多路隔离，自动备份，灾难恢复	去耦合，分段隔离，自动发布订阅	极高，自适应流里均衡

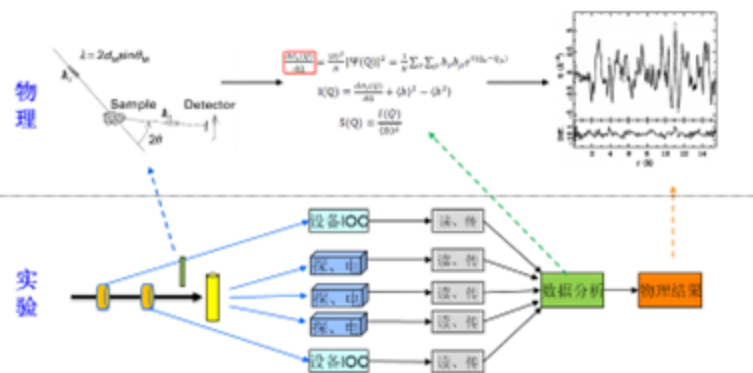
## 数据流标准化



对每一种业务数据流定义明确的接口，实现抽象化的聚合与共享，完善多路数据流的复用和扩展



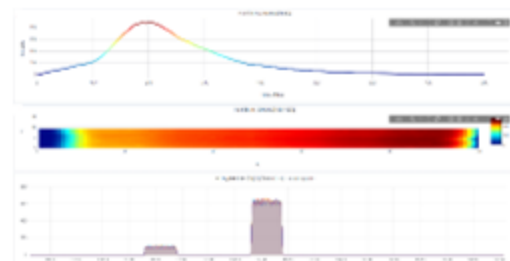
## 分析平台



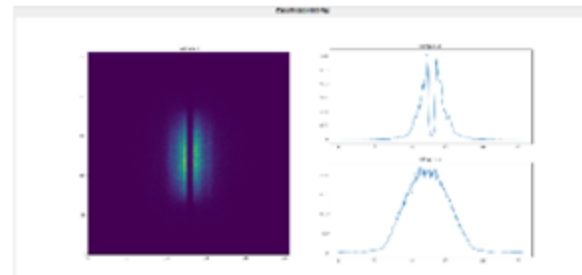
### 从原始数据中处理和提取物理信息

- 提供谱仪联调和实验运行中常用的数据处理模块，软件设施标准化
- 向专业组和谱仪科学家开放环境和共享大数据
- 与控制程序建立反馈通道，对前端数据流进行标记，扩展实验功能

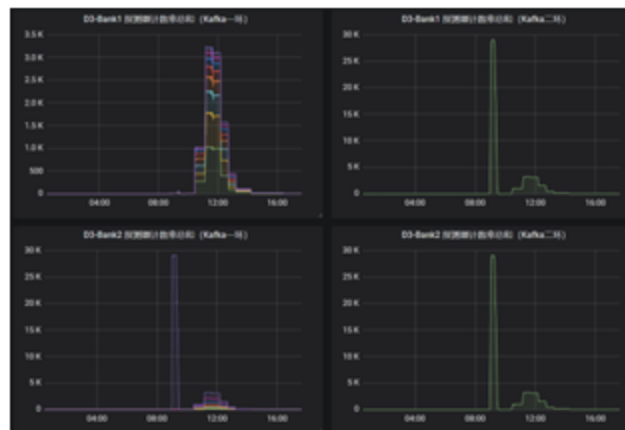
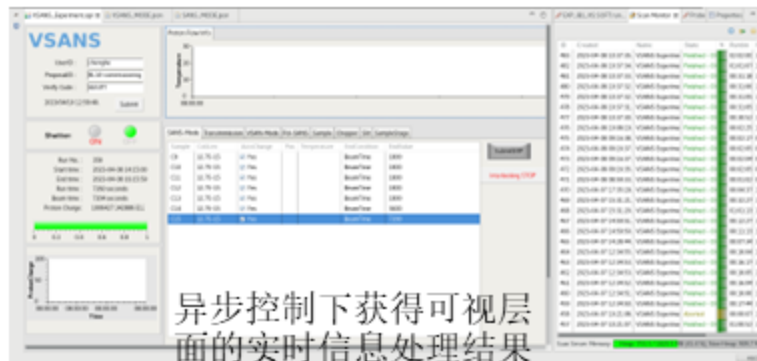
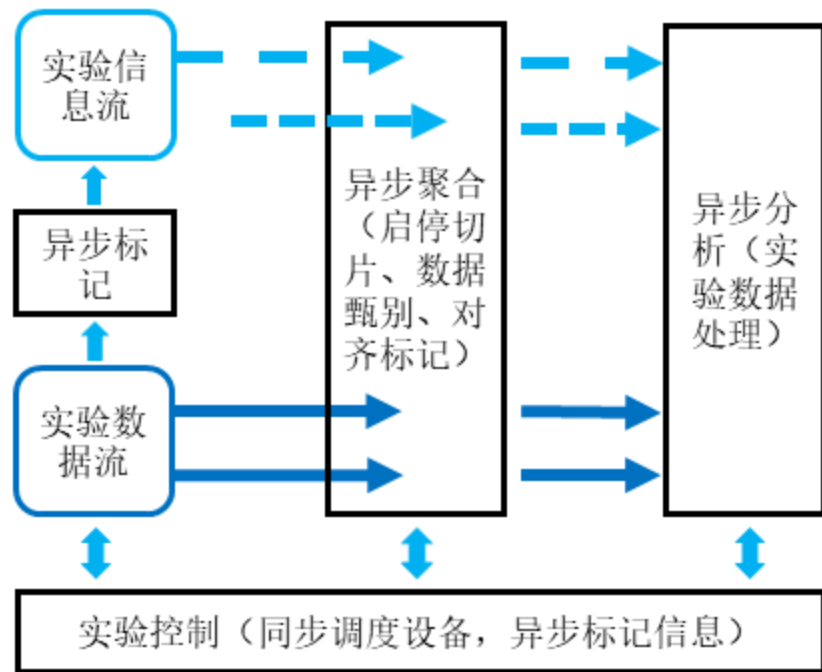
MPI长期监测分析信息



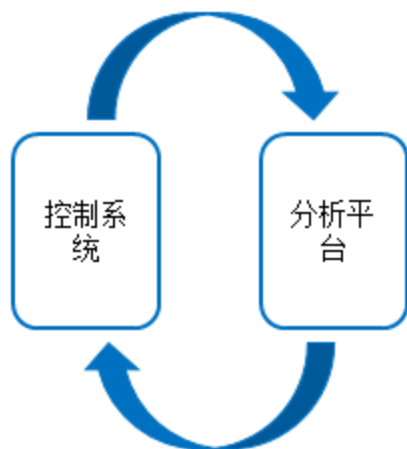
微小角利用分析平台进行BeamStop位置刻度



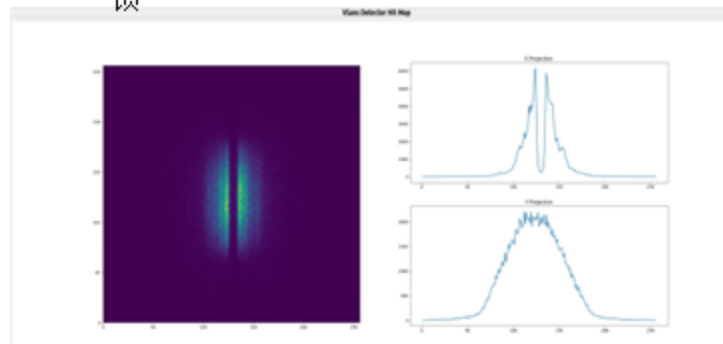
## 控制系统



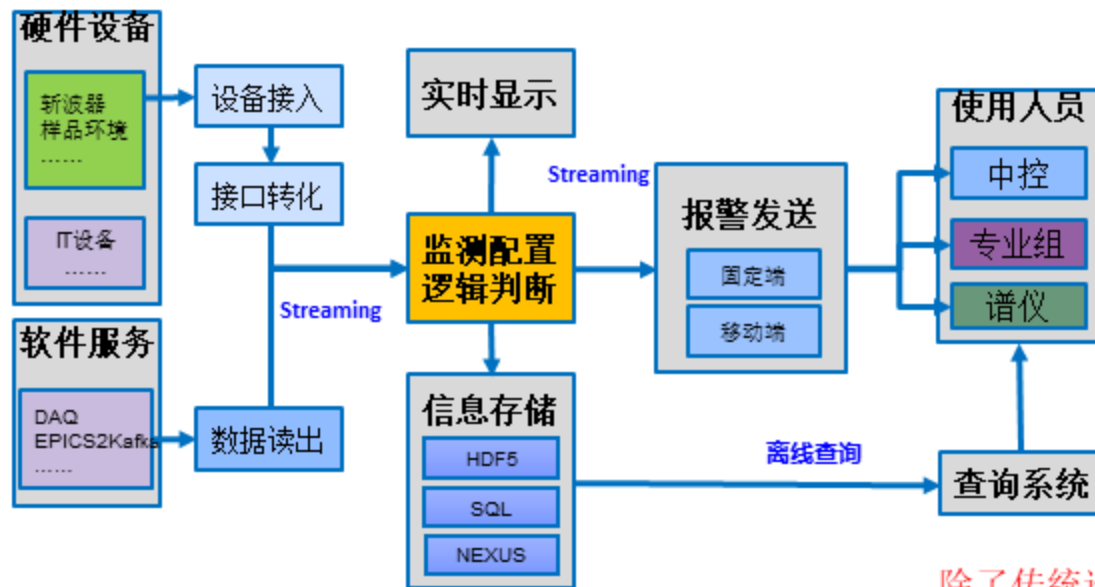
利用分析平台提取的物理信息和可视级实时反馈，实现实验调控的自动/半自动闭环



VSANS-BeamStop刻度分析可视化与反馈



## 监测报警

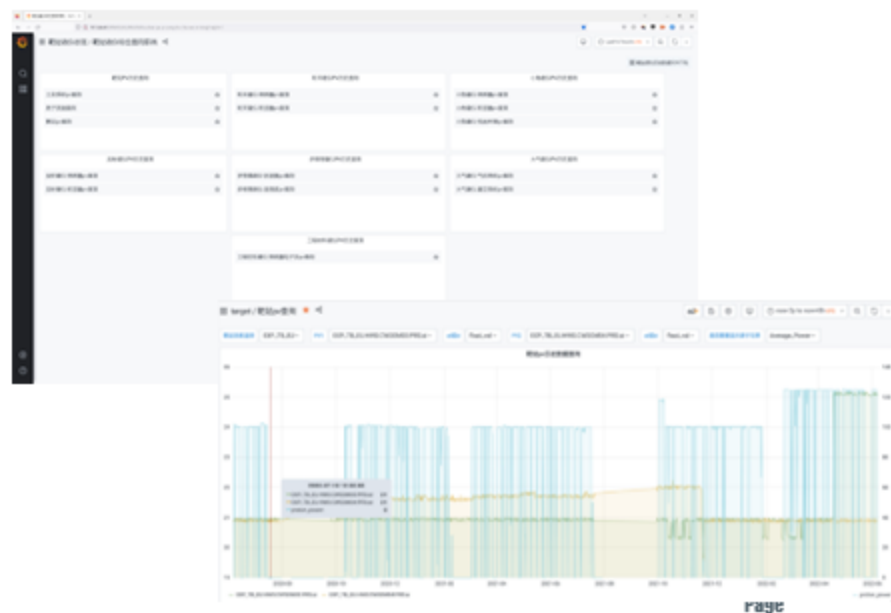


- 利用标准接口采集多种类型信息数据
- 实时监测，向用户开放逻辑
- 分级报警，多路输出，智能判断

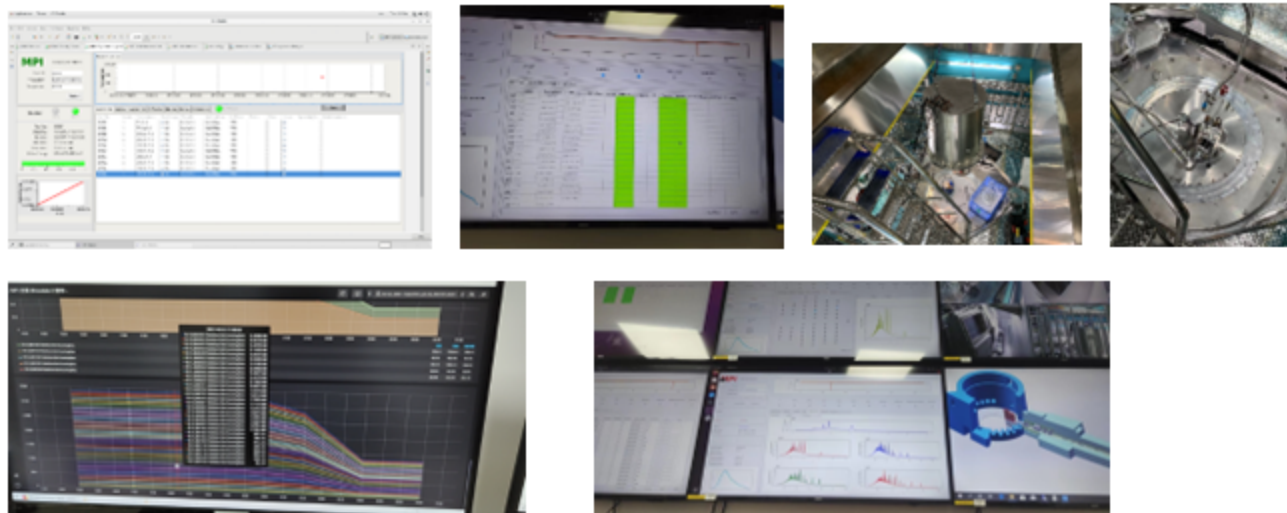
除了传统设备状态报警之外，面向实验状态的高层智能报警功能将是报警系统的下一代升级方向

## 其他周边服务

对接数据流和数据库，可从门户网站导航访问到多种周边服务：  
 历史数据查询和下载，实时监测网页显示，实验报表记录查看，  
 实验控制和运行状态记录（专家界面）。。。



## 成功应用在多条合作谱仪-3运行3调试




多物理谱仪于2021年完成部署DSNI系统，实现基于数据流的准实时数据传输和处理业务。

基于数据标记流的流控子系统和相应的分析平台模块于2021.11进行了部署联调，成功对接online部分和样品环境设备，从此作为多物理的流控方案顺利运行至今。



## 小结和展望



作为衔接先进硬件和新型实验方法的关键，控制和数据处理框架是实现装置高可用的重要环节，对许多实验方法的支持亦有关键和必要意义。

实验方法：多模态聚合、样品倒空间测量、带束联合刻度。。。 (物理目标的必要条件)

逻辑链路：数据流共享、控制反馈、实验编排、实时处理 (高可用的框架设计)

基础硬件：探测器、电子学、WR、高速网络。。。 (高性能的基础)

欢迎各位老师交流合作