



2024年5月-8月工作报告

报告人：张水涵

导师：朱科军

实验物理中心，TDAQ组

时间：2024年08月30日

目录

■ JUNO-SPMT

- 江门现场完成第6,7次避光测试
- 数据格式修改更新数据流和单机版测试系统

■ Taishan Antineutrino Observatory (TAO)

- 混合数据流测试
- 与CD-Waveform & CD-TQ系统的联调测试

■ ROOT-based general online data visualization system (ROBOT)

- 核心功能开发完成
- 性能测试
- 基于机器学习的直方图异常检测模型训练开发

■ 其他工作 & 进度安排

- 两篇小论文投稿 (RDTM & TNS)

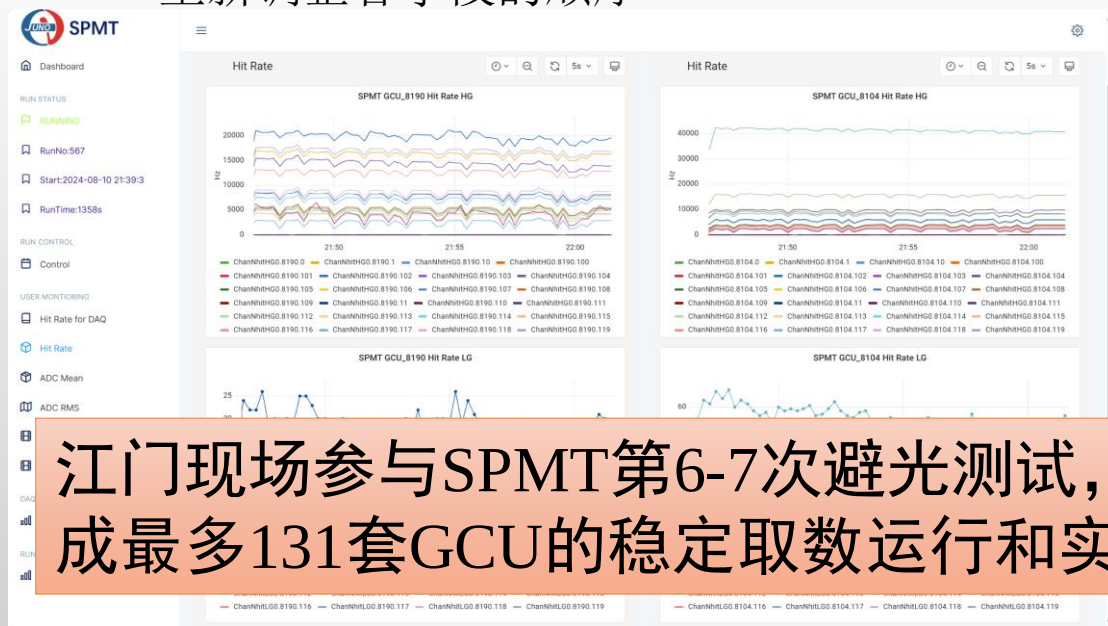
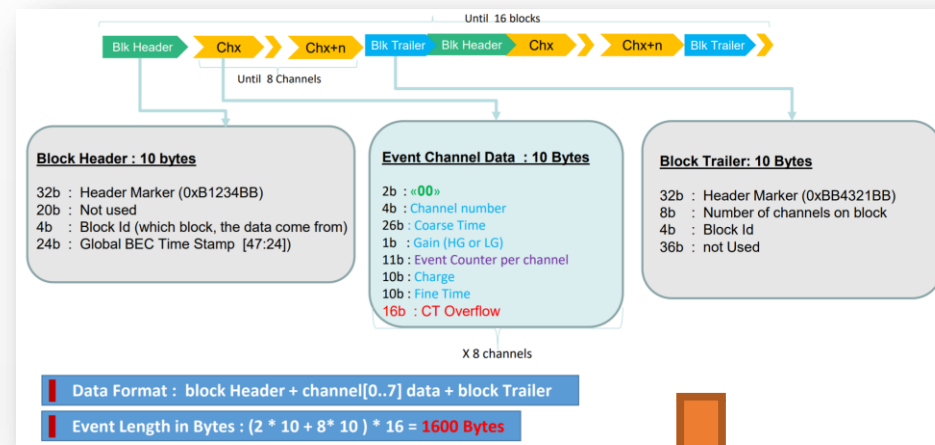
SPMT——江门现场完成两次避光测试

■ 第六/七轮关灯测试（05.26/08.10）——ABCV2.3+GCUV2.0版本固件

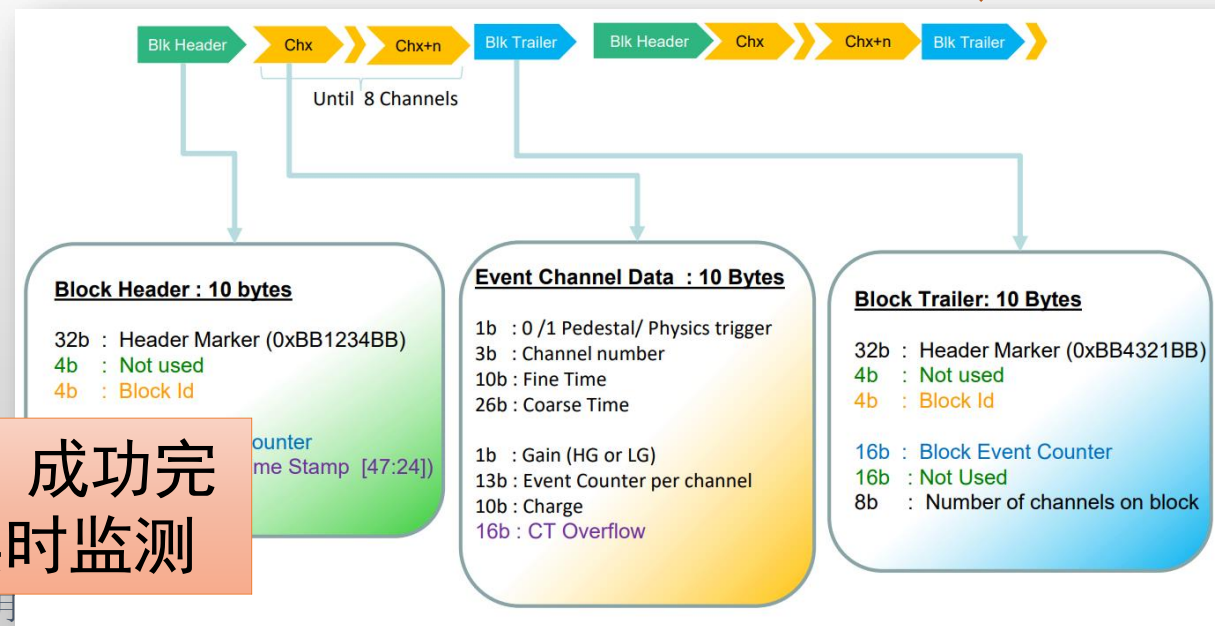
- 电子学配置：IPbus
- 数据读出：TCP协议
- 成功完成131个GCU同时取数运行控制和在线监测

■ 2.41版本固件数据格式升级

- 增加全局时间戳
- 重新调整各字段的顺序



江门现场参与SPMT第6-7次避光测试，成功完成最多131套GCU的稳定取数运行和实时监测



SPMT——测试系统和数据流升级

■ 模拟数据源修改

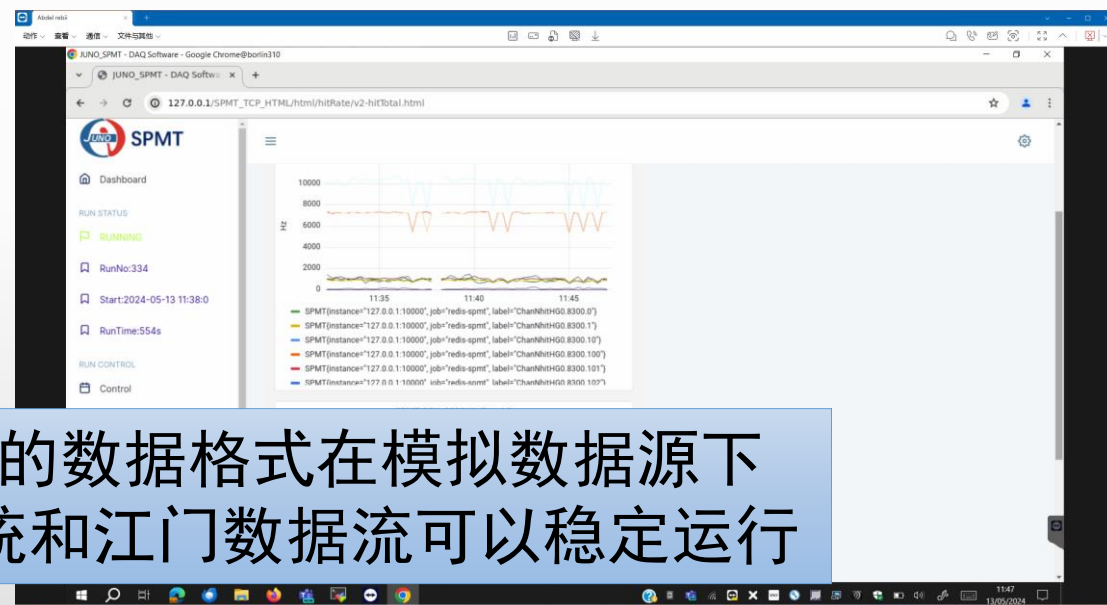
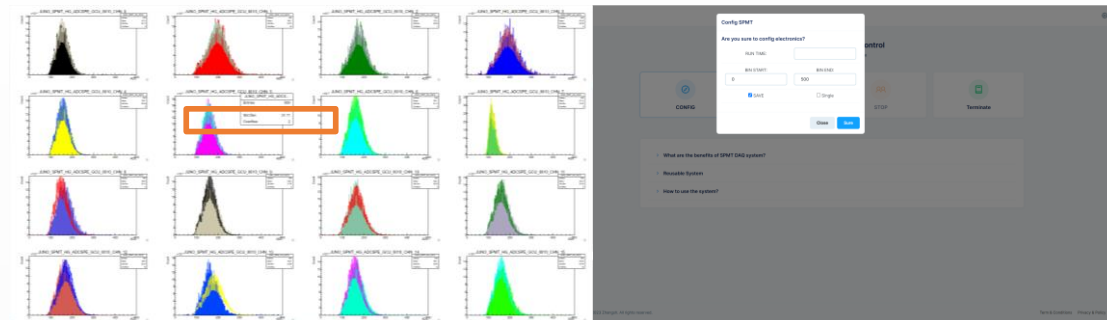
- 原因：江门数据流混合数据测试 & 数据格式改变

■ SPMT DAQ测试系统和数据流升级

- 主要修改数据解析部分
- 使用全局时间戳计算完成时间同步

■ 成功部署法国，应用于实验室的联调测试

- 更通用，便于移植，只需修改setup.sh
- 系统可以稳定运行3天以上



SPMT修改后的数据格式在模拟数据源下
DAQ测试系统和江门数据流可以稳定运行

发现问题：

- 绘制ADC能谱时发现越界，无法展示
- 直方图分bin范围不可配

DAQ软件
升级

解决方案：

- 修改监测部分以及网页增加bin范围可配
- 直方图中状态框中增加overflow展示

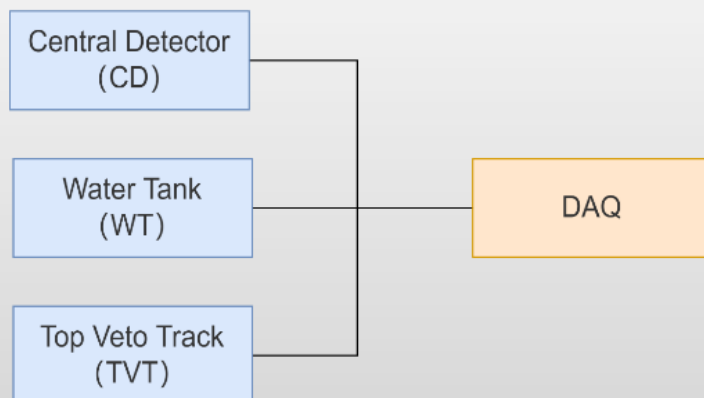
TAO DAQ工作进展

■ 台中微子实验 (Taishan Antineutrino Observatory)

江门中微子实验的子实验

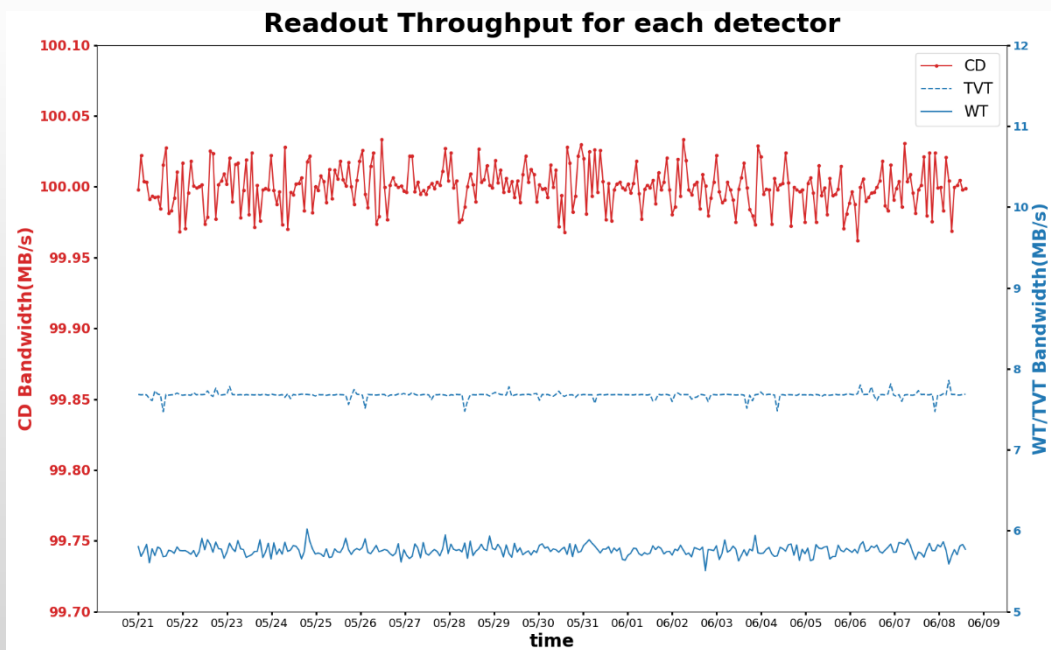
■ TAO DAQ数据流

- 三种探测器对应的数据流分别开发完成
- 验证了在线软件触发算法的正确性
- 为数据流增加在线实时监测
 - 监测系统运行状态



TAO DAQ工作进展——模拟混合数据流测试

- 目的：模拟全规模数据量验证全链路数据流
- 开发模拟数据源为每个子系统
 - CD T/Q, TVT T/Q, WT T/Q data
 - 发送数据带宽和FDR一致
- 最长稳定运行时间：~18天（停电导致）
- 发布了第一个软件版本 **taodaq1.0_juno2.3**



发现问题：

1. 混合数据流只能运行不超过3天ROS挂死
2. DP进行软件触发时，出现大量报错信息，指示数据头不存在
3. DA存在内存泄漏，运行过程内存缓慢上涨



解决方案：

1. 考虑时间戳48b循环时，错误增加了0x80000 0000 0000，实际应该增加0x1 0000 0000 0000
2. 在case语句判断时丢失break，导致一直运行到结束才退出，进入错误的case条件
3. 尚未找到原因

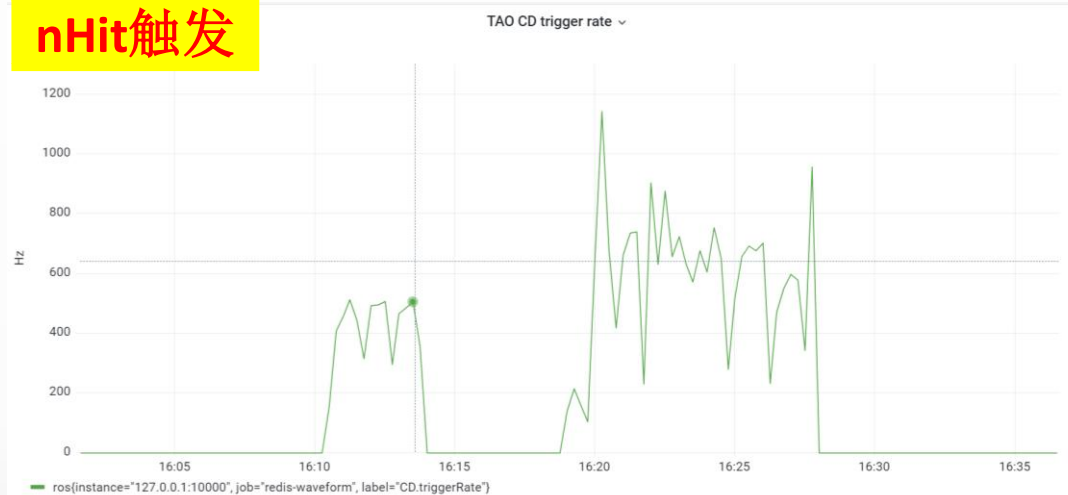
在模拟全规模数据量下验证了全链数据流的稳定性和读出性能

高能所工厂联调测试——SiPM&FEB&FEC&DAQ

CD

- 1: 1小模型实验
- 两种不同的全链路测试系统
- 全链路测试基于IHEP ADC (波形 + TQ)
 - 已经稳定运行 > 6months
 - 共获取 > 5TB数据用于离线分析
- 全链路测试基于意大利ADC (仅TQ数据)
 - ◆ 测试条件: 连接1块CU触发板 & 64通道/FEC
 - ◆ nHit触发
 - 多次联调测试
 - 可以取数时间<10mins (由于FEC数据流时间戳不稳定)
 - ◆ 外触发
 - 可以取数>30mins
 - 触发率相比预期减少一半

nHit触发



外触发



高能所工厂联调测试——塑闪&FEB&FEC&TDAQ

TVT

■ 1: 1小模型实验

■ TVT全链路系统联调测试

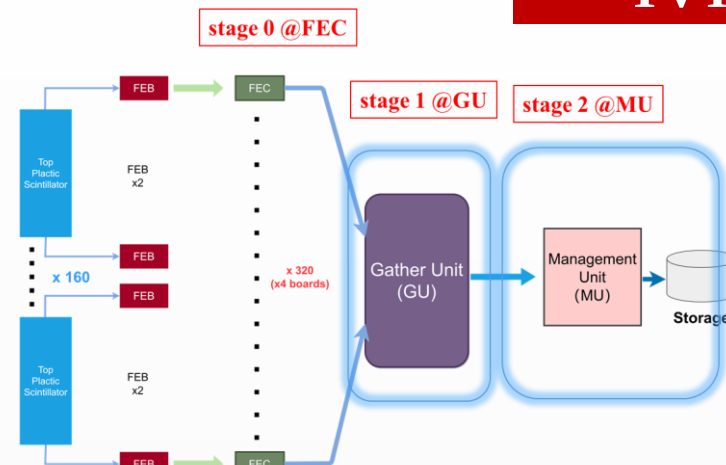
– 塑闪 & FEB & FEC & TGU & DAQ

■ 全链路数据流带有层符合软件触发算法稳定运行

■ AlmaLinux9.4环境安装代码移植

– 最多支持到gcc11，无法手动装gcc8（不支持）

– 测试系统代码移植成功，可以稳定运行



Resolution

Primary compiler packages

- The following gcc packages are available as the default compiler:

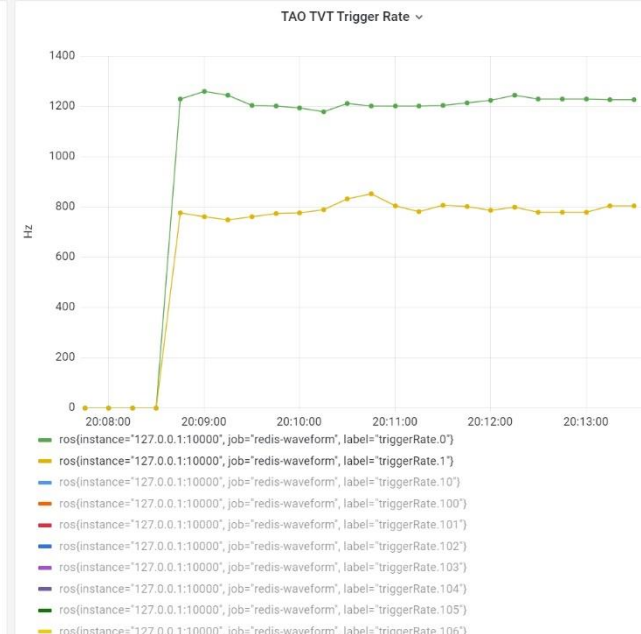
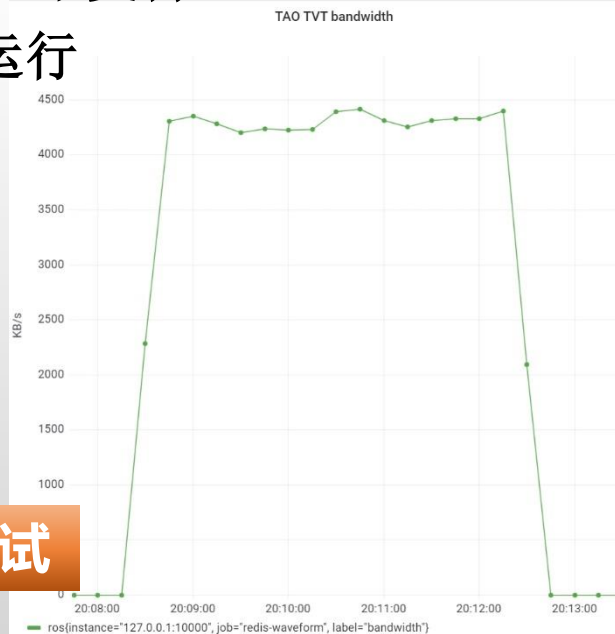
- RHEL9: gcc 11.x in app stream
- RHEL8: gcc 8.x or gcc 9.x in app stream
- RHEL7: gcc 4.8.x
- RHEL6: gcc 4.4.x
- RHEL5: gcc 4.1.x
- RHEL4: gcc 3.4.x
- RHEL3: gcc 3.2.x
- DTS9: gcc 9.1.x
- DTS8: gcc 8.2.x
- DTS7: gcc 7.2.x
- DTS6: gcc 6.2.x
- DTS4: gcc 5.2.x, 5.3.x
- DTS3: gcc 4.9.x
- DTS2: gcc 4.8.x
- DTS1: gcc 4.7.x

Compiler backward compatibility packages

Compatibility packages are available to provide build com

- RHEL9: Red Hat does not have any plan to ship a ba
- RHEL8: Red Hat does not have any plan to ship a ba
- RHEL7:
- compat-gcc-44 (gcc 4.4.7 for compatibility with code designed to be built under RHEL6)
- RHEL6:

准备中山大学继续联调测试



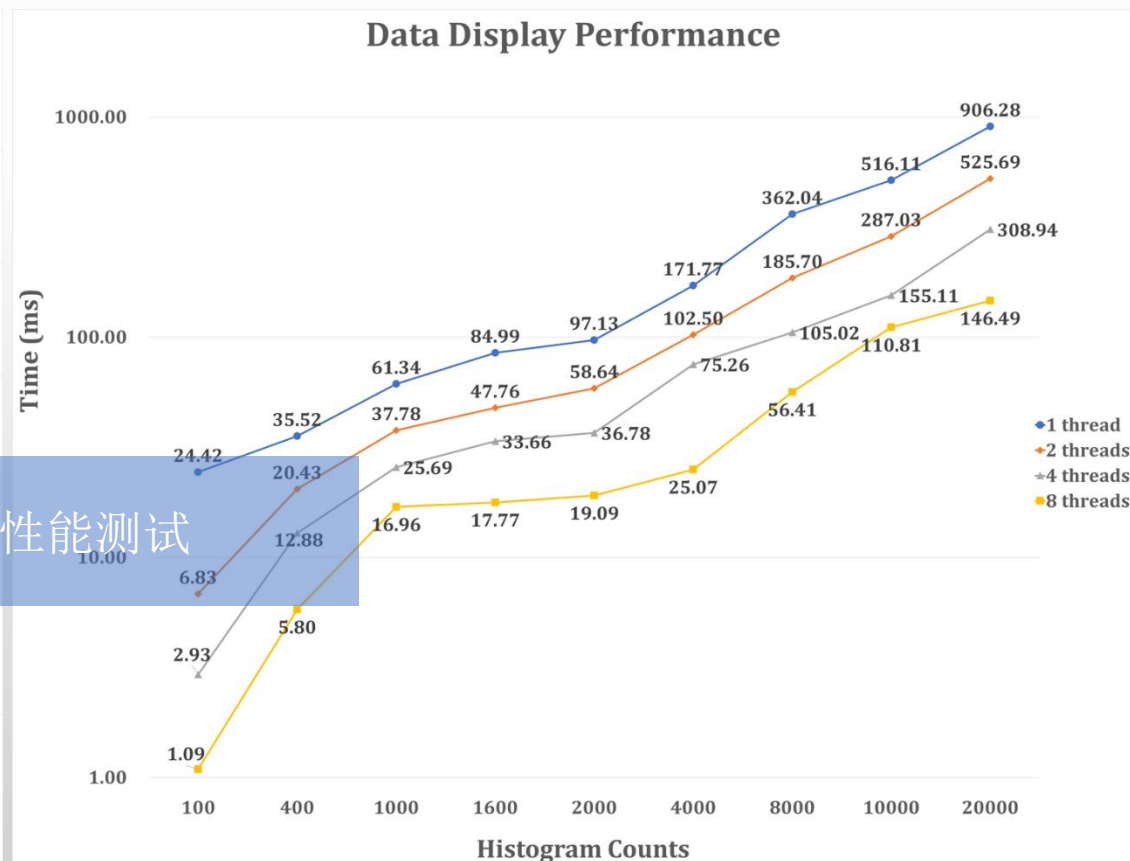
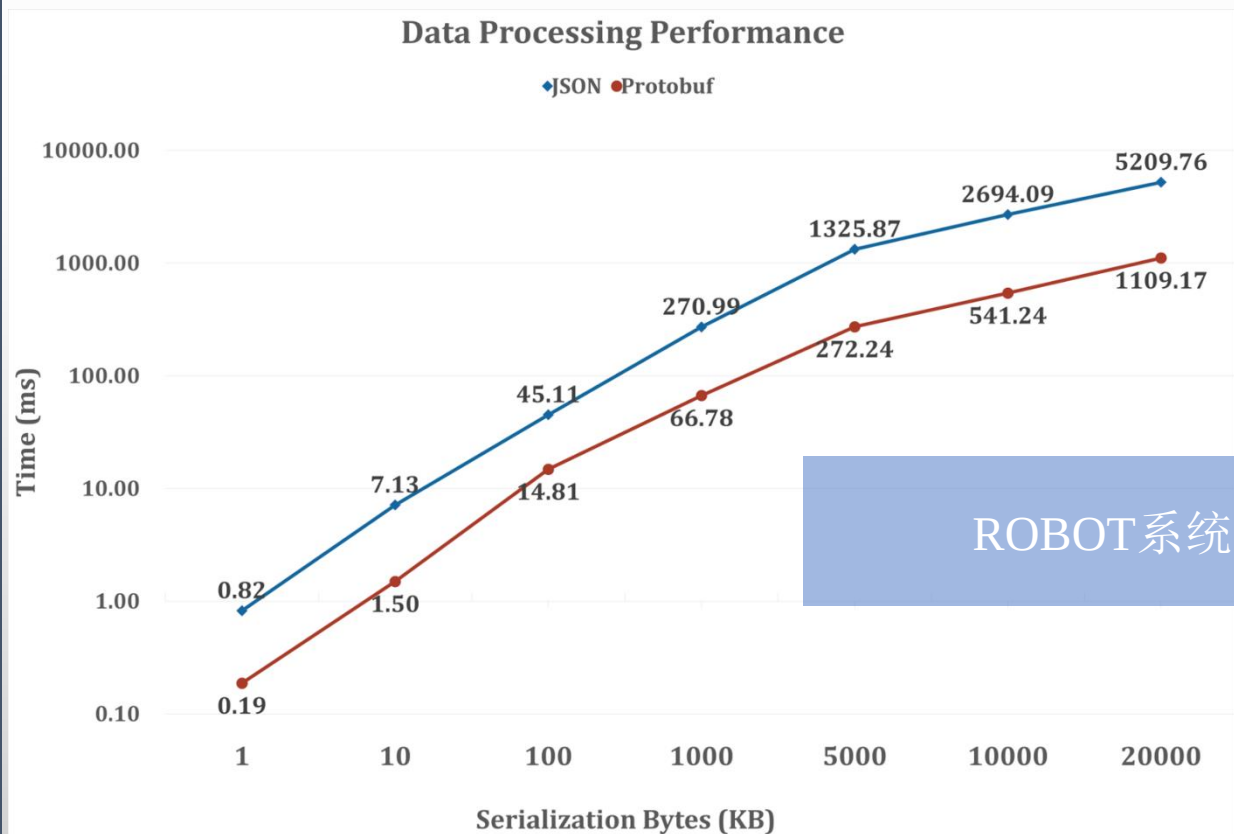
ROBOT——基于ROOT的通用数据可视化系统

■ 目的：通用性和可扩展性，独立于探测器 / 电子学

■ **ROBOT系统优化**

- 画图逻辑重整：只在初始化时画图，后续仅更新图中数据
- 解决增加overflow程序崩溃问题

- 容易被 JUNO + TAO集成
- 已经成功应用于联调测试和江门现场安装SPMT三次关灯测试



ROBOT系统性能测试

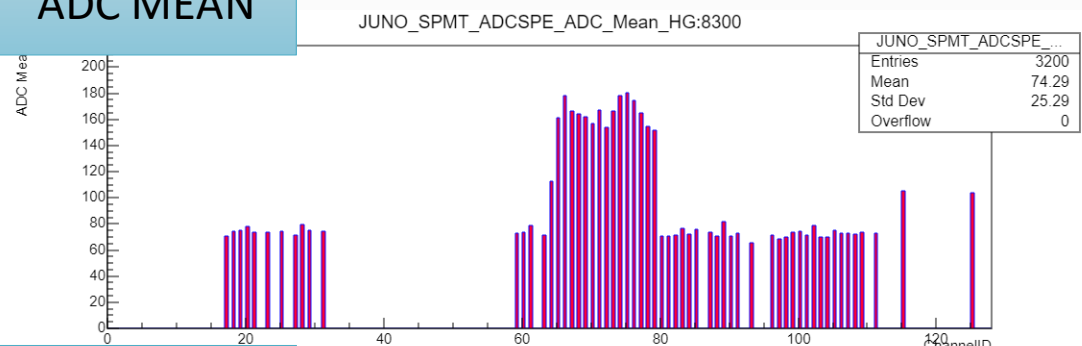
基于机器学习的直方图异常检测

依赖人工智能和机器学习方式实现自动异常检测

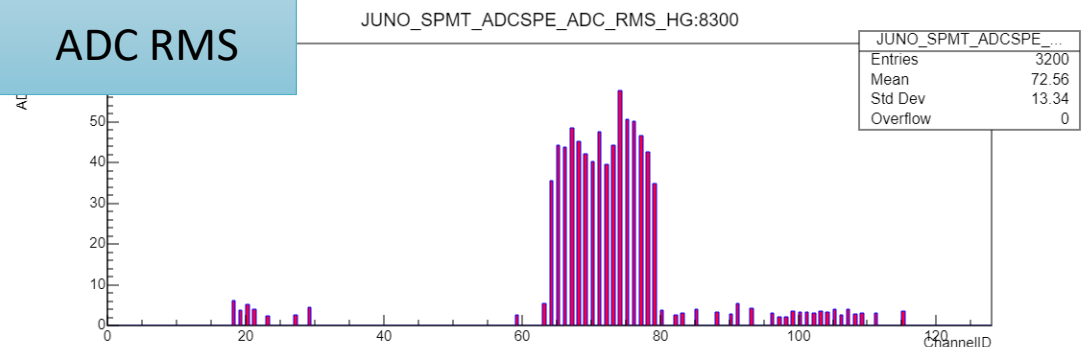
目前高能物理领域异常检测方式

- 基于统计的方式识别低级异常——报警机制
- 数据可视化——直观发现异常数据点或趋势
- 更高级的监控由人工完成——耗时耗力

ADC MEAN



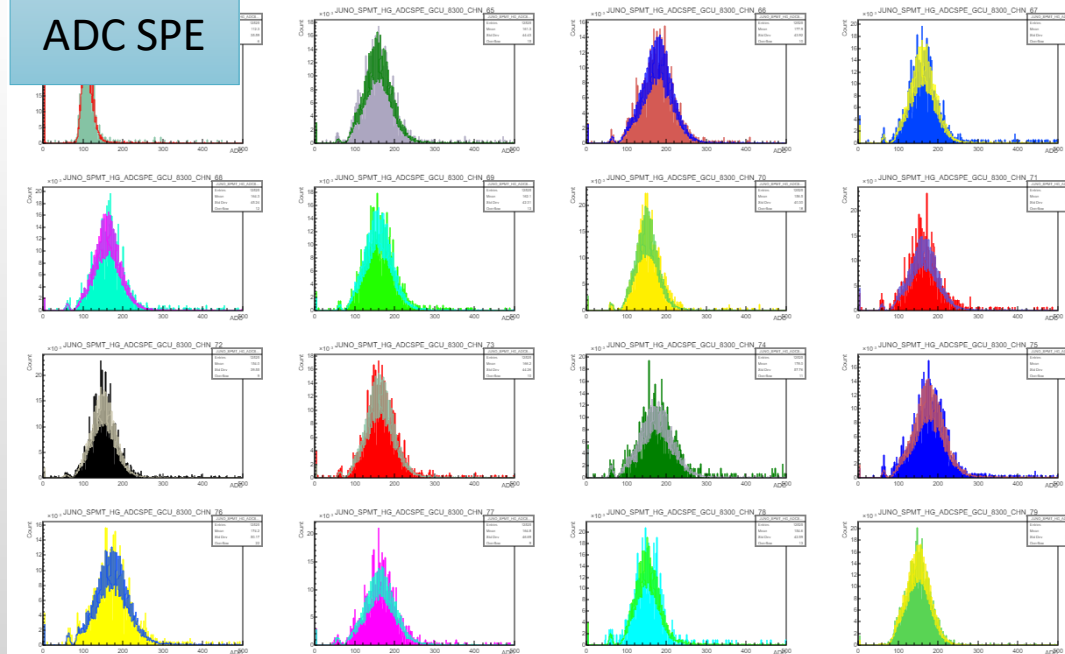
ADC RMS



以JUNO-SPMT异常检测为例（现有方式）

- 统计电荷能谱，均值和方差
- 可视化查看，在线绘制直方图，人工识别异常
- 存在问题：目前关灯测试已经超过~12000个通道，十分耗时

ADC SPE



基于机器学习的直方图异常检测

□ Auto-Encoder (AE), 无监督学习算法

□ 神经网络: 输入层, 隐藏层和输出层

□ 模型主要部分:

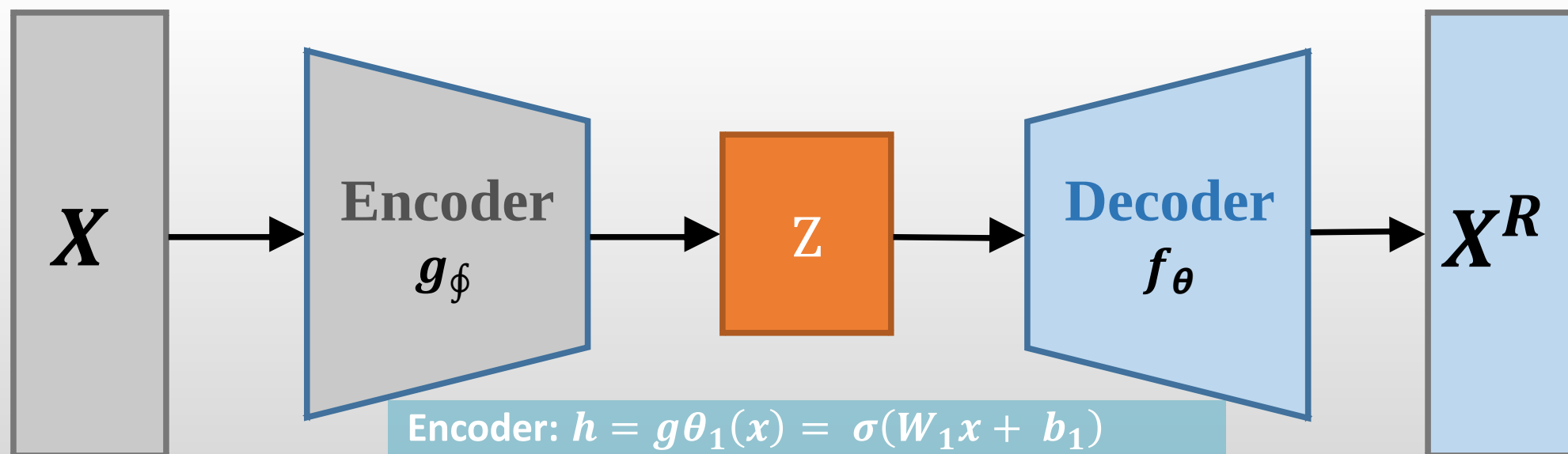
- 编码器: 输入数据压缩成低维表示
- 解码器: 低维表示重构成原始数据维度

□ 输入: SPMT直方图

□ 损失函数: 衡量原始数据与重构数据之间的差异
– 均方误差 (MSE) 或交叉熵损失 (BCE)

□ 训练过程: MLP, 反向传播算法

□ 训练目标: Encoder的输入和Decoder的输出越接近越好



$$\text{Encoder: } h = g\theta_1(x) = \sigma(W_1x + b_1)$$

$$\text{Decoder: } \hat{x} = g\theta_2(h) = \sigma(W_2x + b_2)$$

$$\text{Loss: } \text{dist}(X, X^R)$$

基于机器学习的直方图异常检测

- 使用SPMT第五次和第六次关灯测试中数据 (~20w条样本)
- 每个样本得到损失函数 $l(x)$,按照阈值 λ 来确定异常

$$f(x) = \begin{cases} normal, l(x) < \lambda \\ anomaly, l(x) \geq \lambda \end{cases}$$

- 目标：适当损失精准率的情况下，尽可能提高召回率
- 模型优化：
 - 数据清洗 & 调参（维度，batch size）
 - 损失函数（MSE，KL散度，BCE） & 调整模型结构（三层→六层）
 - Skip connection

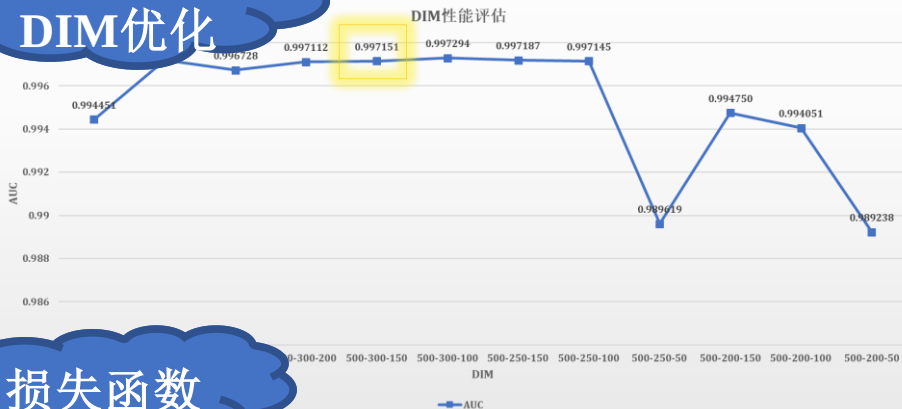
$$\text{召回率} = \frac{\text{预测正确的异常直方图个数}}{\text{实际异常直方图的个数}}$$

$$\text{精准率} = \frac{\text{预测正确的异常直方图个数}}{\text{预测出的异常直方图总数（预测正确+错误）}}$$

SPMT直方图异常检测——模型优化

- 模型调参 (DIM, batch size)
- 损失函数 (MSE, BCE, KL散度)
- 模型结构优化
- Skip connection

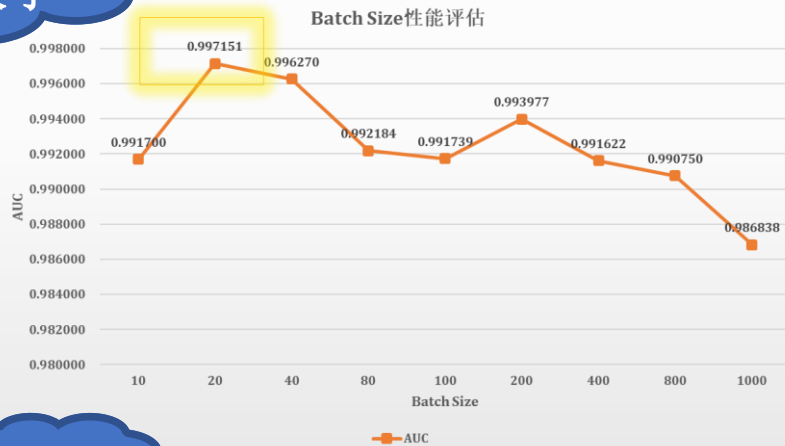
调参—— DIM优化



损失函数

损失函数	AUC	正样本损失值	负样本损失值	总体平均损失值
MSE	0.99715082	11.02346655	0.000550762	1.870566584
BCE	0.948387106	1.305231539	0.01628325	0.234950773
KL散度	0.341959063	482.6341742	483.7929908	483.5963999

调参——批 处理大小



模型结构

层数	AUC	正样本损失值	负样本损失值	总体平均损失值
三层	0.997151	11.023467	0.000551	1.870567
五层	0.971982825	1.41547329	0.011357356	0.249562842

效果最佳:

DIM: 500-300-150

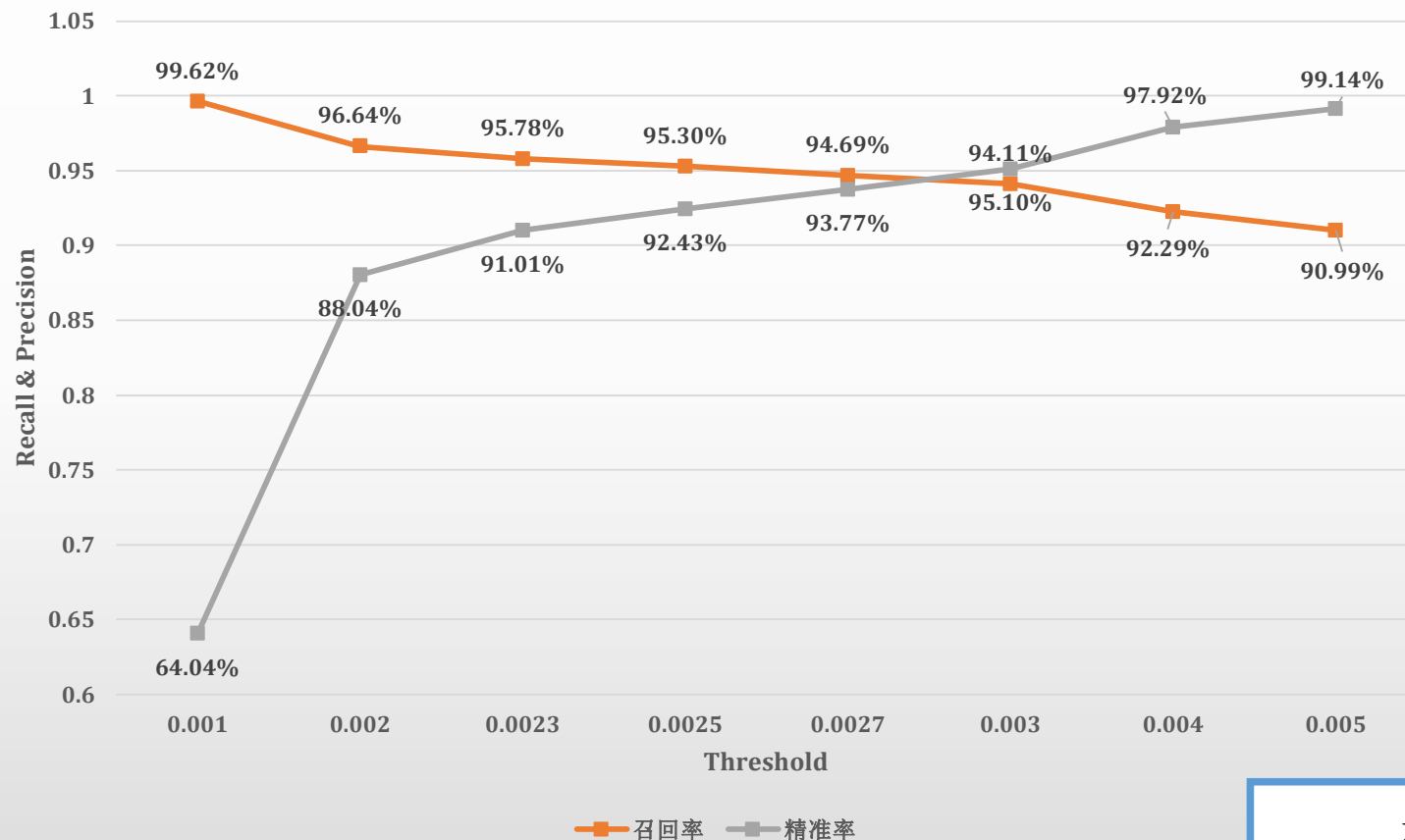
Batch size: 20

Loss: MSE

三层神经网络

SPMT直方图异常检测——模型效果

阈值影响评估



使用模型探测异常时间 (CPU) :

- 加载模型: 0.4348s
- 测试时间: 0.7963s
- 一共检测直方图数量: 62,276条数据样本

$$\text{召回率} = \frac{\text{预测正确的异常直方图个数}}{\text{实际异常直方图的个数}}$$

$$\text{精准率} = \frac{\text{预测正确的异常直方图个数}}{\text{预测出的异常直方图总数 (预测正确+错误)}}$$

SPMT直方图异常检测

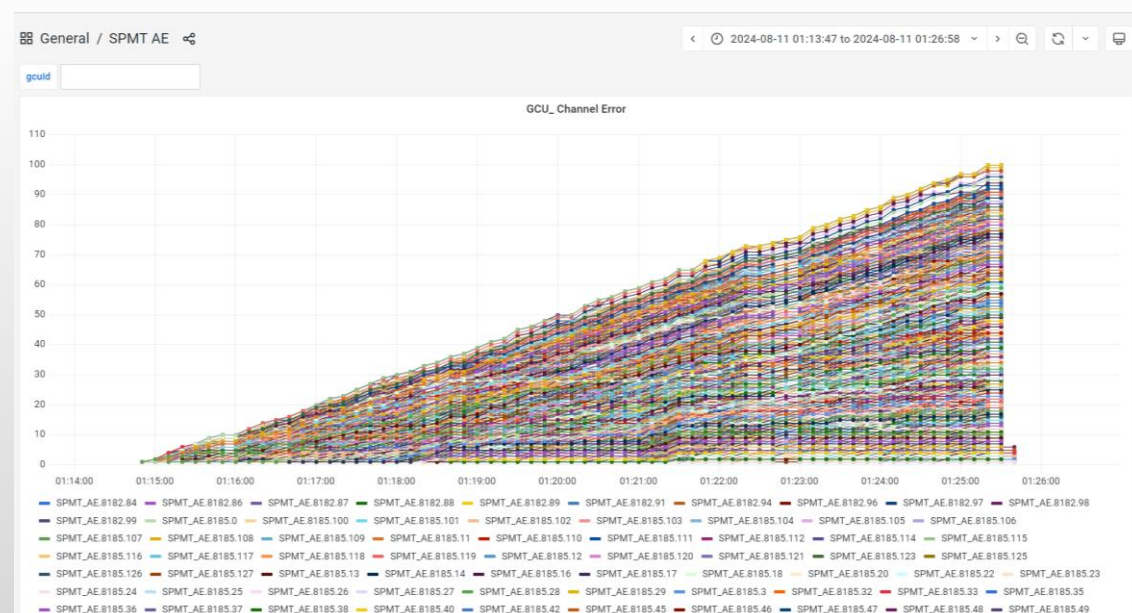
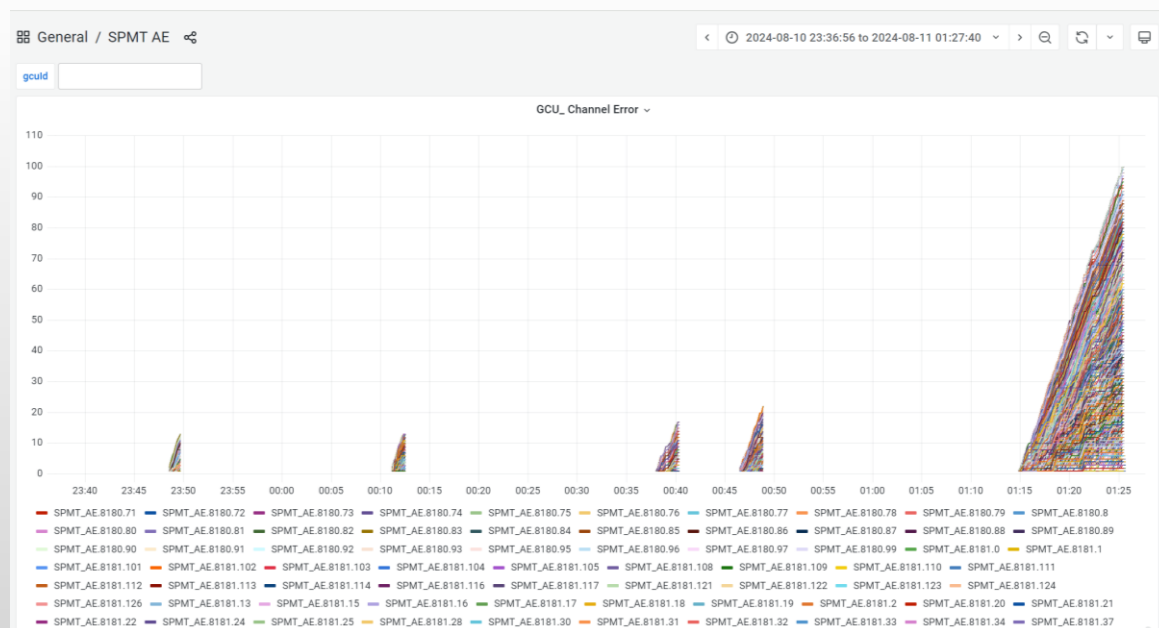
□和数据流接口对接

□全流程跑通

—加载模型，获取数据，过模型，得到异常结果

□成功应用于SPMT第7次关灯测试中

—模型识别出所有人工发现的异常



其他工作 & 进度安排

■ 其他工作

- 参加[24rd江门合作组会](#)并报告TAO TDAQ进展
- TAO DAQ小论文已经投稿RDTM, 小修完成
- ROBOT小论文已经回复意见, 大修正在修改中

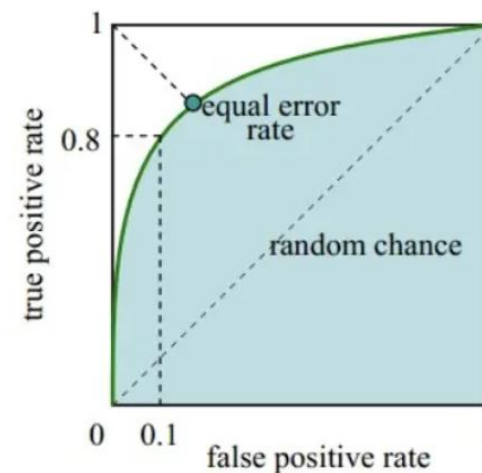
■ 进度安排

- 继续进行TAO CD & TVT全链路系统联调测试
- 增加TAO DAQ数据流直方图监测和Grafana监测, 对接ROBOT系统
- ROBOT系统集成直方图异常检测
- 异常检测模型优化

Thanks!

Backup——SPMT直方图异常检测

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN



- 准确率 = $(TP+TN)/(TP+TN+FP+FN)$
- 精准率 = $TP/(TP+FP)$ (预测出的异常正确度)
- 召回率 = $TP/(TP+FN)$ (预测出异常概率)
- 真正率 (TPR) = 灵敏度 = $TP/(TP+FN)$
- 假正率 (FPR) = $1 - \text{特异度} = FP/(FP+TN)$

评估指标

- ROC (真正率与假正率的关系) —— 分类器在不同阈值下的性能
- AUC (ROC曲线的面积) —— 可分离的程度或测度 (量化评估)
- AUC越高, 模型在将0预测为0, 将1预测为1时越好, 阈值选取